Review

Seismic tomography: A window into deep Earth

N. Rawlinson [a,*], S. Pozgay [a], S. Fishwick [b]

[a] *Research School of Earth Sciences, Australian National University, Mills Rd., Canberra, ACT 0200, Australia*
[b] *Department of Geology, University of Leicester, Leicester LE1 7RH, UK*

A R T I C L E   I N F O

A B S T R A C T

The goal of this paper is to provide an overview of the current state of the art in seismic tomography, and trace its origins from pioneering work in the early 1970s to its present status as the pre-eminent tool for imaging the Earth's interior at a variety of scales. Due to length limitations, we cannot hope to cover every aspect of this diverse topic or include mathematical derivations of the underlying principles; rather, we will provide a largely descriptive coverage of the methodology that is targeted at readers not intimately familiar with the topic. The relative merits of local versus global parameterization, ray tracing versus wavefront tracking, backprojection versus gradient based inversion and synthetic testing versus model covariance are explored. A variety of key application areas are also discussed, including body wave traveltime tomography, surface wave tomography, attenuation tomography and ambient noise tomography. Established and emerging trends, many of which are driven by the ongoing rapid increases in available computing power, will also be examined, including finite frequency tomography, full waveform tomography and joint tomography using multiple datasets. Several practical applications of seismic tomography, including body wave traveltime, attenuation and surface waveform, are presented in order to reinforce prior discussion of theory.

© 2009 Elsevier B.V. All rights reserved.

Contents

* Corresponding author. Tel.: +61 2 6125 5512; fax: +61 2 6257 2737.
  E-mail address: nick@rses.anu.edu.au (N. Rawlinson).

1. Introduction

1.1. What is seismic tomography?

Seismic tomography is a data inference technique that exploits information contained in seismic records to constrain 2D or 3D models of the Earth's interior. It generally requires the solution of a large inverse problem to obtain a heterogeneous seismic model that is consistent with observations. More formally, provided that we can establish an approximate relationship $\mathbf{d} = \mathbf{g}(\mathbf{m})$ between seismic data $\mathbf{d}$ and seismic structure $\mathbf{m}$ – so that for a given model $\mathbf{m}$ we can predict $\mathbf{d}$ – then the seismic tomography problem amounts to finding $\mathbf{m}$ such that $\mathbf{d}$ explains the data observations $\mathbf{d}_{obs}$. In most cases $\mathbf{d}$ and $\mathbf{m}$ are discrete vectors of high dimension, which means that many data records are used to constrain a detailed model. Implicitly, this detail must apply to both vertical and lateral structure. As such, the radial Earth model produced by Backus and Gilbert (1969), based on the theory in their seminal paper of the preceding year (Backus and Gilbert, 1968), is not usually viewed as an early example of seismic tomography despite the similarity in methodology.

A simple example of seismic traveltime tomography, which serves to illustrate several features typical to most applications, is shown in Fig. 1. In this artificial test, a synthetic model in spherical shell coordinates is generated (Fig. 1a) which consists of 780 grid points evenly spaced in latitude and longitude with cubic B-spline functions used to describe a smooth velocity field. For a given set of sources and receivers, first-arriving geometric ray traveltimes are then computed (Fig. 1b) through the model. These traveltimes constitute the synthetic data set that is equivalent to the information one may obtain from seismograms recorded in the field. The relationship $\mathbf{d} = \mathbf{g}(\mathbf{m})$, where $\mathbf{d}$ represents the traveltime dataset and $\mathbf{m}$ the velocity model, is non-linear in this case because the path taken by the seismic energy is a function of velocity. Almost without exception, only the first-arrivals of any phase are exploited in traveltime tomography, because later arrivals due to multi-pathing (wavefront folding) are difficult to pick. One property of first arrivals is that they tend to avoid low velocity anomalies, and preferentially sample high velocity anomalies, as can be see in Fig. 1b.
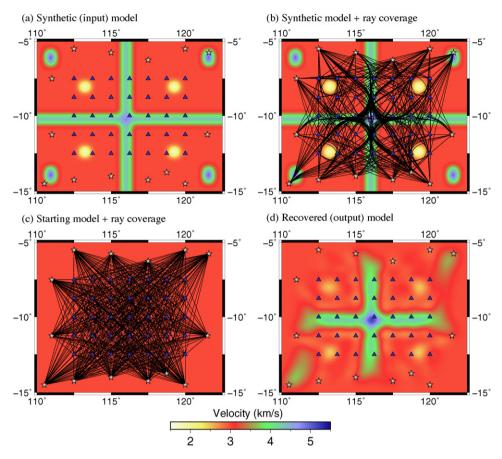
Due to the non-linearity of the inverse problem, the traveltime misfit surface (some measure of the difference between observation and model prediction) may not be a simple smooth function with a well defined minimum. While a fully non-linear solution technique may therefore seem appropriate, the size of the problem usually makes this computationally prohibitive. Instead, some gradient-based technique is often used, which relies on having an initial model "close" to the solution model. Fig. 1c shows the initial model used in this case, which has a uniform velocity, resulting in great circle paths. The node spacing is identical to that of the synthetic model (Fig. 1a), which will favourably bias the recovery of structure. Repeated application of forward ray tracing and linearized inversion eventually produces the solution model shown in Fig. 1d, which remains unchanged with further iterations, and

satisfies the synthetic dataset. This basic approach, which relies on accurate *a priori* information in the form of an initial model, some class of forward solver, and a local inversion technique, is ubiquitous to most forms of tomography, be it traveltime, surface waveform, anisotropy or attenuation.

Comparison of Figs. 1a and d reveals a number of interesting similarities and differences between the synthetic and recovered models. Clearly, regions near the edge of the model that have no path coverage do not deviate from the initial model. In cases where rays exist but have a similar azimuth, recovered anomalies have a tendency to be severely smeared out in the dominant ray path direction (e.g. the two high velocity anomalies in the southwest and northeast corners of the model). Within the bounds of the receiver array, where path coverage is dense, the recovery of structure appears to be accurate, with the exception of the four distinct low velocity anomalies, whose amplitudes are severely underestimated. This problem arises from the fact that first-arrivals avoid low velocity regions, as shown clearly in Fig. 1b, and therefore poorly constrain them. Most forms of seismic tomography, even those that do not directly exploit traveltimes (e.g. attenuation tomography), are affected in some way by this phenomenon, because they usually rely on the paths provided by first-arrival tomography to solve the data prediction problem. Other issues, including solution non-uniqueness (where more than one solution satisfies the data to the same extent) and the validity of geometric ray theory, will be discussed in the following sections.

1.2. Pioneering work

The name most commonly associated with the origins of seismic tomography is that of Keiiti Aki, who published a seminal paper in 1976 on 3D velocity determination beneath California from local earthquakes (Aki and Lee, 1976). In this paper, traveltime data collected at 60 stations from 32 local earthquakes are inverted for 3D crustal structure, described by a total of 264 constant slowness (inverse of velocity) blocks, and hypocenter corrections. The inversion is linear, because ray paths are assumed to be straight, and a damped least squares approach is used to find a solution. Estimates of model covariance and resolution are also made to assess solution robustness A year later, this publication was followed by an equally influential paper which employs teleseismic tomography to image the 3D velocity structure beneath the Norwegian Seismic Array (Norsar) in southeast Norway (Aki et al., 1977). Traveltime residual information from distant (teleseismic) earthquakes is used to constrain structure, which is confined to a local region beneath the array. Constant slowness blocks are again used to describe the lithosphere, although this time, the initial model is defined by constant velocity layers, so ray paths are permitted to bend. However, the inversion is still linear as path geometry is not updated to account for the recovered heterogeneity.

The early work of Aki undoubtedly catalyzed the numerous seismic tomography studies of the crust and lithosphere that soon followed, but a number of other influential developments, arguably

Fig. 1. Synthetic reconstruction test illustrating several typical characteristics of seismic traveltime tomography. (a) Synthetic test model with sources (grey stars) and receivers (blue triangles) superimposed; (b) same model as in (a) but with all first arrival paths plotted; (c) starting model and path coverage for the iterative non-linear inversion; (d) recovered model, which can be compared with (a). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

of similar importance, occurred at around the same time. In global tomography, Adam Dziewonski published a paper in 1977 that uses nearly 700,000 P wave travel time residuals from the bulletins of the International Seismological Centre (ISC) to image the velocity structure of the Earth's mantle, described using a spherical harmonic parameterization (Dziewonski et al., 1977). Despite the size of the traveltime dataset, the number of unknowns in the inverse problem is restricted to only 150, presumably due to limitations in computing power.

Although not as frequently cited as Aki and Dziewonski's seminal works, an earlier paper by Bois et al. (1972) implements a scheme that clearly conforms to the above definition of seismic tomography. In this study, the authors use cross-hole (or well to well) active source seismic imaging to examine part of the Lacq oil field of southwest France. Small charges were inserted down one hole, and their detonation recorded by geophones placed down another hole. Traveltimes picked from the resultant seismograms are then inverted for the 2D velocity structure of the cross-section separating the two boreholes. Rather than use a constant block parameterization, a regular grid of nodes is specified together with an interpolant that ensures continuity of the velocity field and its first derivative at every point. The traveltime prediction problem is solved using a shooting method of ray tracing that fully accounts for isotropic heterogeneity, and an iterative non-linear approach, similar to that demonstrated in Fig. 1, is used to reconcile observed and model traveltimes (Bois et al., 1971). Although the number of unknowns that are solved for is 110, and the maximum number of ray paths used is 90, the proposed technique is sophisticated, particularly considering the minimal development

that had occurred in the field prior to this application. One might argue that seismic tomography implies 3D imaging, but in terms of the underlying theory, there is no real difference, except for the size of the inverse problem, and the complexity of the forward solver.

1.3. The last three decades: a brief history

1.3.1. Local studies of the crust and upper mantle

Following the pioneering efforts in seismic tomography described above, a veritable cascade of new applications and developments soon followed. In cross-hole tomography, various techniques for ray tracing and inversion were trialled (McMechan, 1983, 1987; Bregman et al., 1989), but essentially, the underlying method of Bois et al. (1971) was not significantly advanced upon. Backprojection inversion techniques, inherited from medical imaging, were generally more popular than gradient-based methods, perhaps due to similarities in acquisition geometry. Greater innovation came in the form of diffraction and wave equation tomography (Pratt and Worthington, 1988; Pratt and Goulty, 1991; Song et al., 1995; Pratt and Shipp, 1999), which attempt to exploit more of the recorded waveform. Other classes of seismic tomography that have their origins in exploration include reflection tomography and wide-angle (refraction and wide-angle reflection) tomography, which use artificial sources such as explosions, airguns and vibroseis to generate seismic energy. Reflection tomography is a natural compliment to migration imaging, because it offers a means to constrain velocity and interface depth using traveltimes and, less commonly, geometric spreading amplitudes

and reflection/transmission coefficients. One of the first studies to implement reflection tomography was that of Bishop et al. (1985), which combines ray shooting and a gradient based inversion technique to constrain a 2D model described by constant velocity gradient blocks and cubic spline interfaces. Similar studies have also been carried out by Farra and Madariaga (1988) and Williamson (1990). In general, coincident reflection traveltime data alone appears to be insufficient to satisfactorily resolve the trade-off between interface depth and layer velocity. Consequently, more recent efforts have been directed towards joint inversion of traveltime and amplitudes (Wang and Pratt, 1997; Wang et al., 2000), joint inversion of reflection and wide-angle traveltimes (Wang and Braile, 1996; McCaughey and Singh, 1997) and full waveform tomography (Hicks and Pratt, 2001; de Hoop et al., 2006).

Wide-angle tomography is similar to reflection tomography, but the source–receiver offset tends to be much greater in order to detect refracted rays from significant depths (e.g. Pn waves from the Moho). Both 2D and 3D experiments are common, and over the last few decades have played a major role in unravelling the crustal architecture of continents and margins in various parts of the world including Canada (e.g. Hole, 1992; Kanasewich et al., 1994; Clowes et al., 1995; Zelt and White, 1995; Morozov et al., 1998; Zelt et al., 2001, 2006) and Europe (e.g. Riahi and Juhlin, 1994; Staples et al., 1997; Darbyshire et al., 1998; Louden and Fan, 1998; Mjelde et al., 1998; Korenaga et al., 2000; Morgan et al., 2000; Bleibinhaus and Gebrande, 2006). Early efforts in this field tended to treat the wide-angle reflection and refraction data separately, but it was soon recognised (e.g. Kanasewich and Chiu, 1985) that their joint inversion dramatically increased the likelihood of resolving both interface structure and velocity variation. In recent years, wide-angle tomography has been the subject of much interest in the emerging field of full waveform tomography, where the prospect of far greater resolution has motivated a number of different studies (Pratt et al., 1996; Sirgue and Pratt, 2004; Brenders and Pratt, 2007; Jaiswal et al., 2008).

Following the early work of Aki and Lee (1976), local earthquake tomography (or LET) has become a popular tool for imaging subsurface structure in seismogenic regions. One distinguishing feature of the technique is the need to relocate hypocenters in tandem with recovering seismic structure. Although the conceptual basis of LET has not really changed since Aki's original paper, several advances have been made, including full 3D ray tracing and iterative non-linear inversion (Eberhart-Phillips, 1990); direct inversion for V_P/V_S or Q_P/Q_S ratio (e.g. Walck, 1988); development of methods for constraining 3D anisotropic velocity variations (Hirahara, 1988; Eberhart-Phillips and Henderson, 2004) and attenuation structure (Sanders, 1993; Tsumura et al., 2000); and double difference tomography (Zhang and Thurber, 2003; Monteiller and Got, 2005), which aims to significantly improve hypocenter relocation. In subduction zone settings, recent advances include tomographic inversion of shear wave splitting measurements for anisotropic fabric (e.g. Abt and Fischer, 2008), and of velocity and attenuation anomalies for water content, temperature and composition (Shito et al., 2006).

Teleseismic tomography has been used extensively to map the structure of the crust and lithosphere in 3D (e.g. Oncescu et al., 1984; Humphreys and Clayton, 1990; Benz et al., 1992; Glahn and Granet, 1993; Achauer, 1994; Saltzer and Humphreys, 1997; Graeber et al., 2002; Rawlinson et al., 2006b; Rawlinson and Kennett, 2008). Compared to the original technique of Aki et al. (1977), most teleseismic tomography now uses iterative non-linear inversion coupled with 3D ray racing or wavefront tracking (e.g. VanDecar and Snieder, 1994; Steck et al., 1998; Rawlinson et al., 2006b). In most cases, teleseismic tomography is still based on the recovery of isotropic velocity models from arrival time residu-

als, although attempts have been made to recover anisotropy (e.g. Plomerová et al., 2008).

Detailed local studies of the upper mantle have also been conducted using data from surface waves. For regions with close spacing of broadband seismometers, interstation measurements or array techniques (Friederich and Wielandt, 1995; Forsyth and Li, 2005; Pedersen et al., 2003) can be used to estimate the local dispersion characteristics within the zone of interest. These methods have been applied in a number of locations to produce detailed tomographic images of the lithospheric mantle (e.g. Weeraratne et al., 2003; Bruneton et al., 2004; Li and Burke, 2006; Darbyshire and Lebedev, 2009).

1.3.2. Regional and global tomography

The different classes of seismic tomography discussed above tend to use temporary deployments of recorders to target a limited geographical region; hence they can be described as "local" methods. By contrast, regional and global tomography studies more commonly utilize information from permanent networks that span large continental regions or much of the globe, such as the GSN (Global Seismic Network), in addition to any available data from temporary arrays. Targets include the upper mantle, whole mantle or the entire Earth. Since the pioneering work of Dziewonski et al. (1977), which used the traveltimes of P-waves, efforts have been focused on improving resolution by exploiting an ever increasing volume of recorded data. Current global P-wave mantle models that exploit traveltime data from the ISC commonly constrain structure at a scale length of a few 100 km or less using millions of paths (Zhao, 2004; Burdick et al., 2008).

In addition to direct P-waves, other phases such as PcP and PKP are now commonly used to improve coverage, particularly in the core (Vasco and Johnson, 1998; Boschi and Dziewonski, 2000; Karason and van der Hilst, 2001). While spherical harmonics are still preferred in some cases, most body wave studies now opt for local parameterizations, such as blocks or grids, which are better suited for recovering detailed structures such as mantle plumes or subducting slabs (van der Hilst et al., 1997; Bijwaard et al., 1998; Karason and van der Hilst, 2001; Zhao, 2004). The highly uneven data coverage that typifies regional and global body wave studies – due largely to irregular distribution of earthquakes and recording stations – has stimulated the idea of using irregular parameterizations, where blocks or nodes are placed only where they are required by the data. Bijwaard et al. (1998), Bijwaard and Spakman (2000) and Spakman and Bijwaard (2001) use a spatially variable cell size parameterization based on ray sampling, in which an underlying regular grid is used to construct a mosaic of non-overlapping irregular cells. Sambridge and Gudmundsson (1998) propose a more general scheme based on Delaunay and Voronoi cells, which is subsequently applied to whole Earth tomography (Sambridge and Faletic, 2003; Sambridge and Rawlinson, 2005).

Body wave tomography using S-waves is also common in regional and global studies (e.g. Grand et al., 1997; Vasco and Johnson, 1998; Widiyantoro et al., 2002), and can either be done in isolation or simultaneously with P-waves to obtain V_P/V_S ratio as in LET. An alternative is to jointly resolve bulk sound and shear velocity (Su and Dziewonski, 1997; Kennett, 1998), quantities that can be more readily linked to experimental laboratory measurements of the physical properties of mantle minerals.

Surface waves and normal modes can also be used to construct tomographic images of the Earth's interior. Compared to body waves, surface waves have the advantage that they can sample the upper mantle beneath ocean basins at sufficient density to produce well constrained models of oceanic lithosphere; on the other hand, they cannot probe into the deep mantle at high resolution, and have difficultly resolving crustal structure. A variety of different method-

ologies have been applied to obtain information from the surface wavetrain. Some global studies use long paths and attempt to measure phase velocity directly for the fundamental mode for each path (e.g. Ekström et al., 1997; Laske and Masters, 1996). Group velocities can be extracted using filter analysis, and have been used to produce maps at both regional (e.g. Ritzwoller and Levshin, 1998; Danesi and Morelli, 2000; Pasyanos and Nyblade, 2007), and global scales (Shapiro and Ritzwoller, 2002). Additional information from surface wave overtones provides better resolution with depth; with this in mind, van Heijst and Woodhouse (1997) develop a new method for measuring overtone phase velocities. Combining these data with information from body waves, Ritsema et al. (2004) produce a shear wavespeed model of the mantle, with particular emphasis on the upper mantle transition zone. Alternatively, an inversion procedure can be used to fit the surface waveform (and in some cases long period S-waves). This style of approach has also been used to produce shear wavespeed models at both regional (e.g. van der Lee and Nolet, 1997; Simons et al., 1999; Friederich, 2003; Heintz et al., 2005; Fishwick et al., 2005; Priestley et al., 2008) and global scales (Debayle et al., 2005; Lebedev and van der Hilst, 2008).

Normal modes or free oscillations of the Earth, which can be viewed as very long period standing surface waves, also offer a means to constrain seismic structure. Individual peaks of the discrete spectrum are often split due to Earth rotation, ellipticity and lateral heterogeneity. Isolating the latter effect enables both mantle structure (Li et al., 1991; Resovsky and Ritzwoller, 1999), and core structure (Ishii and Tromp, 2004) to be imaged. The advantage of this approach is that data coverage is relatively uniform, but due to the very low frequencies of detectable normal modes, the scale length of recovered heterogeneity tends to be extremely broad.

Another area of active research in global seismology is attenuation tomography, in which lateral variations in the anelastic parameter Q are retrieved. A key challenge with this technique is to successfully extract the anelastic signal from the recorded waveform, which is dominated by elastic effects. Studies to date tend to use surface waves and hence focus on the upper mantle (Romanowicz, 1995; Billien and Lèvêque, 2000; Selby and Woodhouse, 2002; Gung and Romanowicz, 2004; Dalton and Ekström, 2006; Dalton et al., 2008), although body wave studies have also been carried out (Bhattacharyya et al., 1996; Reid et al., 2001; Warren and Shearer, 2002). One of the attractions of attenuation tomography is its strong sensitivity to temperature variations, and therefore its potential to image hot spots, mantle plumes and subduction zones.

Anisotropy is a potentially complex issue in all tomographic studies from local to global scales, as it pervades many regions of the Earth including the crust, upper mantle, core–mantle boundary and inner core. The main barrier to its accurate recovery in tomography is the under-determined nature of the inverse problem; it is difficult enough to resolve isotropic velocity variations, let alone all 21 independent elastic constants required to describe arbitrary anisotropic anomalies. As a result, studies that attempt to include anisotropy do so with a limited subset of the elastic moduli. One of the first studies to resolve upper mantle transverse isotropy with a vertical axis of symmetry – otherwise known as radial anisotropy (requiring five independent parameters) – was that of Nataf et al. (1984), who inverted both Love and Rayleigh wave data for velocity structure, described by degree 6 spherical harmonics, to a depth of approximately 450 km. By assuming this class of anisotropy, downwelling and upwelling features associated with slab subduction were successfully imaged. Radial anisotropy is now frequently incorporated into global shear velocity studies (e.g. Panning and Romanowicz, 2006; Kustowski et al., 2008). A form of anisotropy that is more commonly assumed in surface wave tomography studies is that of azimuthal anisotropy (e.g. transverse isotropy with a horizon-

tal axis of symmetry), which allows velocity to vary as a function of horizontal direction, and is therefore more well tuned to upper mantle dynamics associated with contemporary plate tectonics. Early work in this area was carried out by Tanimoto and Anderson (1984, 1985), who found variations of anisotropy in the upper mantle to be as large as 1.5%, albeit with low order spherical harmonics. Montagner and Nataf (1986) and Montagner and Tanimoto (1991) develop a scheme which they describe as "vectorial" tomography, which allows radial and azimuthal anisotropy to be simultaneously constrained by inversion of surface waveforms and regionalization of phase or group dispersion curves. The incorporation of seismic anisotropy in one form or another in surface wave tomography has now become almost routine (e.g. Debayle, 1999; Simons et al., 2002; Debayle et al., 2005; Sebai et al., 2006), but issues still remain as to the appropriate choice of elastic parameters, and how they may trade-off in an intrinsically under-determined inverse problem.

Shear wave splitting provides insight into the strength and orientation of anisotropy by measuring the differential arrival time between orthogonal components of an arriving shear wave. However, due to the path integral nature of the measurements, it provides limited information on the spatial distribution of anisotropy. In the last few years, this limitation has been addressed in the form of shear wave splitting tomography (e.g. Zhang et al., 2007; Abt and Fischer, 2008), which attempts to map the anisotropy inferred from the splitting measurements into a volumetric model. In related developments, splitting intensity measurements from SKS waves (Favier and Chevrot, 2003) have also been used to perform anisotropy tomography (Chevrot, 2006; Long et al., 2008).

1.4. Recent trends: ambient noise and finite frequency tomography

Recordings of identifiable wavetrains from sources such as earthquakes or explosions form the basis of traditional methods of seismic tomography as described above. However, since the turn of the millennium, virtual-source seismology has gradually emerged to become an important field in modern seismology, thanks to the work of a number of researchers who have both theoretically and experimentally demonstrated a remarkable property of fully diffuse or random wavefields: information they accumulate about the medium through which they propagate can be extracted by the long-term cross-correlation of waveforms recorded at two separate locations (e.g. Lobkis and Weaver, 2001; Campillo and Paul, 2003; Shapiro and Campillo, 2004; Snieder, 2004; Wapenaar et al., 2005; Sabra et al., 2005; Wapenaar and Fokkema, 2006). It turns out that the cross-correlation produces an estimate of the Green's function between two points; that is, the signal that would arrive at one point if the source waveform were a delta function (or point impulse) located at the other point. This is a particularly useful piece of information, because the travel time and shape of the wavelet are purely a function of the properties of the intervening medium. In the seismic case, the cross-correlation of ambient seismic noise recorded at two stations (Shapiro and Campillo, 2004), or the seismic coda associated with distant earthquakes (Campillo and Paul, 2003), can be used to extract empirical Green's functions. For the seismic coda, multiple scattering from small-scale heterogeneity in the lithosphere appears to generate a sufficiently diffuse wavefield. Oceanic and atmospheric disturbances, further randomized by scattering caused by solid Earth heterogeneity, is one of the main energy sources for ambient noise tomography.
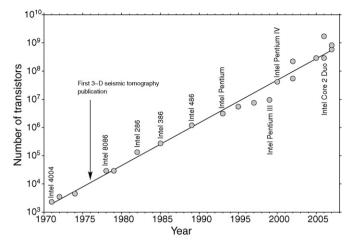
Ambient noise tomography has now become an established technique for imaging Earth structure at a variety of scales, but its development continues at a rapid pace. The most common approach is to extract Rayleigh wave group traveltimes from the cross-correlated waveforms and invert for group velocity at different periods (e.g. Shapiro et al., 2005; Sabra et al., 2005; Kang and

Shin, 2006; Yao et al., 2006; Yang et al., 2007; Zheng et al., 2008). This is often done under the assumption of straight/great circle ray paths, but several studies have used bent rays via solution of the eikonal equation, and have therefore addressed the non-linearity of the problem (Rawlinson et al., 2008; Saygin and Kennett, 2009). In the seminal study of Shapiro et al. (2005), only one month of data from the US Array stations was required to produce high resolution images of the California crust, which clearly discriminates between regions of thick sedimentary cover and crystalline basement. More recent efforts have been directed towards recovering phase velocity in addition to group velocity (Benson et al., 2008), and attempting to resolve 3D shear wave velocity structure from the inversion of Rayleigh and Love wave dispersion maps (Benson et al., 2009).

Geometric ray theory forms the basis of the forward prediction problem in most forms of seismic tomography, but its validity is limited to cases where the seismic wavelength is much smaller than the scale length of heterogeneity that characterizes the medium through which it passes. In fact, unless the seismic energy has infinitely high frequency (which of course is unphysical), the actual ground motion recorded by a seismometer will have a partial dependence on the medium in the neighbourhood of the geometric ray. Unless properly accounted for, this finite frequency effect will essentially blur the final image. Recognition of this fact has been longstanding in the seismic imaging community, but until recently, a workable solution was impeded by limits in both computing power and theoretical development. One of the first surface wave studies that attempted to account for finite frequency effects was that of Snieder (1988a,b), who used so-called first-order perturbation theory (or Born theory) to account for scattering. The new technique was used in the inversion of waveform phase and amplitude to construct phase velocity maps of Europe and the Mediterranean.

In the context of body wave tomography, sensitivity kernels for traveltimes or waveforms have been formulated by a variety of researchers (e.g. Luo and Shuster, 1991; Yomogida, 1992; Vasco and Mayer, 1993; Li and Romanowicz, 1995; Friederich, 1999; Marquering et al., 1999; Dahlen et al., 2000; Zhao et al., 2000). The intriguing result that body wave traveltimes are insensitive to heterogeneity exactly along the geometric ray path led Marquering to use the terminology "banana doughnut" kernel. Using such sensitivity kernels, finite frequency body wave tomography has been applied to a number of different datasets with often interesting results (e.g. Montelli et al., 2004; Yang et al., 2009), not least of which are the well defined mantle plumes revealed in the study of Montelli et al. (2004).

It is briefly worth noting that the beginnings of finite frequency tomography were accompanied by some discussion as to its validity in general heterogeneous media and the degree of improvement it brought to conventional ray-based tomography (de Hoop and van der Hilst, 2005a,b; Dahlen and Nolet, 2005; Montelli et al., 2006; Trampert and Spetzler, 2006). However, with increasing use of the technique, and validation against wave equation solvers (Tromp et al., 2005), these discussions have become less relevant. Besides the study of Montelli et al. (2004), others to have used finite frequency tomography include Hung et al. (2004), who report increased resolution in the upper mantle transition zone beneath Iceland; Chevrot and Zhao (2007), who use finite frequency Rayleigh wave tomography to image the Kaapval craton; and Sigloch et al. (2008), who exploit teleseismic P-waves to elucidate the structure of subducted plates beneath western North America.

Compared to seismic traveltime tomography based on geometric ray theory, the advantage of finite frequency traveltime tomography is that a larger range of phase information is used to constrain structure. For a single source–receiver arrival, filtering over a large range of frequencies will produce a set of delay

Fig. 2. Increase in the number of transistors as a function of time for a range of Intel microprocessors. [Source: 60 years of the Transistor: 1947–2007, Intel website.]

times (e.g. extracted using cross-correlation with a synthetic pulse - see Nolet, 2008) that can be inverted for structure. The advantage of phase information is that it behaves more linearly than the waveform, and is hence more amenable to inversion by linearized techniques. Another benefit of finite frequency tomography is that it is feasible to invert amplitude information (e.g. Sigloch et al., 2008) due to the phenomenon of wavefront healing. Geometric ray amplitudes behave in a much more non-linear fashion, and are therefore difficult to incorporate in tomography. Provided that broadband observables are available, finite frequency tomography has the potential to improve seismic imaging on many fronts.

1.5. Seismic tomography and computing power

The rise of seismic tomography is inextricably linked with the rapid advances in digital computing and microprocessor technology that began in the 1960s. This branch of seismology would simply not be feasible without the ability to make millions to trillions of calculations per second. An often used proxy for the growth in computing power is Moore's law, which stems from his seminal paper (Moore, 1965) in which he predicted that the number of components on an integrated circuit would increase exponentially, approximately doubling every two years up until at least 1975. Today, this rule of thumb is applied to the number of transistors on a microprocessor. Fig. 2 shows the "Moore's law" plot for Intel processors between 1971 and 2007. Linear regression demonstrates that an exponential increase appears to be a valid approximation, at least until recently. With recent emphasis on cluster computing and multiple core processors, the rapid increases in computing power appear set to continue.

Early applications of seismic tomography were challenged by what we would probably today regard as breathtakingly limited hardware. For example, at about the time that Keiiti Aki's pioneering work on seismic tomography was published (Aki and Lee, 1976; Aki et al., 1977), the cutting edge in computing power was the Cray I, the first commercially successful vector computer. It was capable of a peak performance of 250 million FLOPS (floating point operations per second)—though usually ran at about 80 million FLOPS, had about 8 megabytes of main memory, and weighed nearly 2.5 tons (Schefter, 1979). Although impressive for its time, the stunning advances in computing power over the last few decades means that a standard desktop computer is now many times faster. For example, computers using a single Intel Core i7 processor can expect performance of around 60+ billion FLOPS (source: www.hardcoreware.net), some 750 times faster than the

CRAY 1 in normal operating mode. While this increase in power is considerable, it should be considered in the context of the vast quantities of high quality digital seismic data that are now being recorded and archived, and the need to represent the Earth by many parameters to properly accommodate such large quantities of information in tomographic studies. For example, recent regional and global body wave tomography studies use over 10^7 traveltimes to constrain models with 10^5–10^6 unknowns (e.g. Burdick et al., 2008).

The remainder of this review paper will describe methods used in seismic tomography for representing structure, solving the forward and inverse problems, and assessing solution non-uniqueness. Several case studies of local and regional tomography are then presented to provide the reader with a broad cross-section of the different types of studies that are commonly carried out, and the particular issues associated with them. The final section of the paper will discuss future directions in seismic tomography. Within the confines of a relatively short review paper, it is not possible to cover every aspect of this large and diverse field. In addition to the many references that are provided in specific subject areas, we can recommend several other review articles and books. The edited volumes of Nolet (1987) and Iyer and Hirahara (1993) are notable for being two of the earliest books to be published on the subject and contain a wealth of useful information. However, they do not provide a gentle introduction to the subject. The recent book by Nolet (2008) provides an authoritative, coherent and wide ranging dissertation on seismic tomography and is highly recommended. The review article of Rawlinson and Sambridge (2003b) provides good coverage of crustal and lithospheric traveltime tomography, and the review paper of Romanowicz (2003) is a good introduction to global mantle tomography.

2. Representation of structure

Ideally, one would like to extract structural information from seismic data without first imposing limitations on the nature of its spatial variation. In the synthetic example shown in Fig. 1, cubic B-splines on a regular grid were used to represent velocity structure, which limited the minimum scale length of heterogeneity to the chosen grid spacing, and only allowed smooth variations in wavespeed. In reality, the Earth may contain both continuous and discontinuous (e.g. Moho, faults) variations in wavespeed, and exhibit structural heterogeneity at multiple scale-lengths. Hence, our choice of parameterization immediately restricts the field of permissible models, and can be viewed as a form of ad hoc regularization. The use of splines on a regular grid to represent structure, as in Fig. 1, is an example of a regular static parameterization, which is by far the most common approach used in seismic tomography. Other options include irregular parameterizations, where the minimum scale length of structure is variable, and adaptive parameterizations, where the inversion process plays a role in adjusting the number and/or location of parameters to suit the resolving power of the data. While several studies have used static irregular parameterizations, they are generally applied within an adaptive framework. Apart from limiting the range of structure that can be recovered, the choice of parameterization is important because it impacts on the solution technique chosen for both the forward and inverse problems.

2.1. Common regular parameterizations

Regular parameterizations are attractive because they are conceptually simple, easy to formulate, and generally do not complicate the forward and inverse solvers. Cells or blocks (Fig. 3a) with uniform seismic properties (e.g. velocity or slowness) are the most basic form of parameterization, and make initial value ray tracing simple because path segments in each block are straight lines. On the other hand, the artificial discontinuities between each block are unrealistic, and can lead to unwarranted ray shadow zones and triplications, which may make the two-point ray tracing problem more non-linear. Using a large number of blocks with some form of smoothing regularization can mitigate these problems, but it will be at the expense of increased computing time. Constant slowness/velocity blocks have been widely used in most forms of tomography, including teleseismic (Aki et al., 1977; Oncescu et al., 1984; Humphreys and Clayton, 1988, 1990; Benz et al., 1992; Achauer, 1994; Saltzer and Humphreys, 1997), local earthquake (Aki and Lee, 1976; Nakanishi, 1985), wide-angle (Zhu and Ebel, 1994; Hildebrand et al., 1989; Williamson, 1990; Blundell, 1993) and global (Grand et al., 1997; Vasco and Johnson, 1998; van der Hilst et al., 1997; Boschi and Dziewonski, 1999). A slightly more sophisticated approach is to use triangular cells (2D) or tetrahedra (3D) with a constant velocity gradient, which like constant velocity blocks, facilitates analytic ray tracing (e.g. Chapman and Drummond, 1982; White, 1989).

An alternative to block parameterizations is to define seismic properties at the vertices of a regular grid (Fig. 3b) together with some interpolation function. One of the first implementations of this approach was by Thurber (1983), who used trilinear interpolation between a rectangular grid of nodes to define a continuously varying velocity field for local earthquake tomography. This scheme is now commonly used in earthquake tomography (Eberhart-Phillips, 1986, 1990; Zhao et al., 1992; Eberhart-Phillips and Michael, 1993; Scott et al., 1994; Graeber and Asch, 1999), and can be found in other forms of tomography, including teleseismic tomography (Zhao et al., 1994; Steck et al., 1998). The use of higher order interpolation results in a smoother continuum, but requires a larger basis. For example, trilinear interpolation means that any point within a cell is defined by the 8 points that describe the cell, but produces C_0 continuity (i.e. continuous, but not differentiable everywhere). On the other hand, the use of cubic B-splines means that any point within a cell is a function of 64 surrounding points, but results in C_2 continuity (i.e. continuous second derivatives). Thus, there is generally a trade-off between smoothness, the width of the local basis, and consequently, computing time. Exceptions include natural cubic splines, which are cubic polynomials that interpolate each grid point and have a global basis (i.e. any point defined by the spline is a function of all grid points). Cubic spline functions with a local basis are used widely in tomography: Thomson and Gubbins (1982) and Sambridge (1990) use Cardinal splines in teleseismic and local earthquake tomography respectively; Farra and Madariaga (1988) and McCaughey and Singh (1997) use cubic B-splines in wide-angle tomography; and Rawlinson et al. (2006b) use cubic B-splines in teleseismic tomography. Splines under tension (Smith and Wessel, 1990) is a flexible form of parameterization that essentially allows variation between quasi-trilinear interpolation and cubic spline interpolation. The ideal tension factor results in a smooth model that minimizes unrealistic oscillations yet maximizes local control. Neele et al. (1993), VanDecar et al. (1995) and Ritsema et al. (1998) all use this approach in teleseismic tomography.

In regional and global tomography, regular blocks or grids in spherical coordinates are faced with the additional challenge of an artificial increase in spatial resolution towards the poles and central axis. To address this problem Wang and Dahlen (1995) and Wang et al. (1998) develop spherical surface splines which essentially correspond to a cubic B-spline basis on a triangular grid of approximately equally spaced knot points. In global waveform tomography, the so-called "cubed-sphere" (Ronchi et al., 1996), which is an analytic mapping from the cube to the sphere, has become popular, particularly in conjunction with the spectral element method for
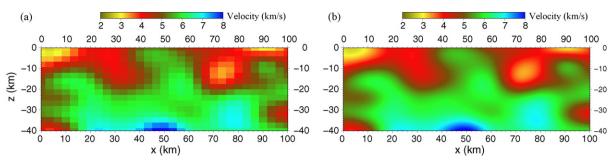
Fig. 3. 2D velocity field defined using (a) constant velocity blocks; (b) cubic B-spline patches.

numerical solution of the elastic wave equation (Komatitsch et al., 2002).
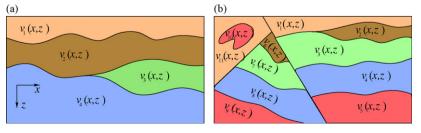
A common alternative to the discretization of seismic properties in the spatial domain is to instead use the wavenumber domain. Spectral parameterizations that use some form of truncated Fourier series fall into this category (e.g. Wang and Houseman, 1997); the unknown parameters in the inversion problem then become the amplitude coefficients of the harmonic terms, rather than the values at grid nodes or within blocks that is generally the case when the spatial domain is parameterized. At the local scale, spectral parameterizations have been used in wide-angle traveltime inversion by Hildebrand et al. (1989), Hammer et al. (1994) and Wiggins et al. (1996). In global tomography, spherical harmonics are frequently used for structural representation (Dziewonski et al., 1977; Dziewonski and Woodhouse, 1987; Li et al., 1991; Trampert and Woodhouse, 1995; Reid et al., 2001; Romanowicz and Gung, 2002) due to their natural affinity with the shape of the Earth, their relative simplicity in controlling the wavelength of recovered structure, and their common usage in other global geophysical studies (e.g. gravity, magnetism), which helps facilitate direct comparison. The drawback of infinitely differentiable functions of this type is that they have a global basis (i.e. adjustment of any single harmonic term will have a global influence on the model), so poorly resolved portions of a model may detrimentally influence (or "leak") into other regions. Furthermore, compute time can become significant for models described by a large number of harmonic terms, since they all contribute to the value of the function at any given point. Amirbekyan et al. (2008) attempt to address these shortcomings by developing a harmonic spherical spline parameterization, which combines spherical harmonics with the spatial localization properties of spline functions.

Representing the Earth by a function which assumes continuous variation of seismic properties is valid in many circumstances, but there are cases where explicit inclusion of interfaces is required. For example, in wide-angle tomography, refracted and reflected waves are the primary observables, and cannot be synthesized without the presence of discontinuities (one could argue that sharp velocity gradients will give rise to similar phenomena, but the data will generally not be able to resolve such features, so explicit interfaces are a valid approximation). There are two basic styles of interface

parameterization that are used in seismic tomography. The most common represents the subsurface as one or more sub-horizontal layers overlying a half-space (Fig. 4a); each layer laterally traverses the entire model, but may pinch together in one or more places (Rawlinson and Sambridge, 2003b). This is often used in coincident reflection and wide-angle tomography, where ubiquitous interfaces such as the Moho are well suited to this form of representation (e.g. Chiu et al., 1986; Farra and Madariaga, 1988; Williamson, 1990; Sambridge, 1990; Wang and Houseman, 1994; Zelt, 1999; Rawlinson et al., 2001a; Rawlinson and Urvoy, 2006). The velocity (or other seismic property) within each layer can be represented using any of the techniques described above, and need not necessarily be linked to the interface geometry or adjacent layers. The relative simplicity of this representation makes it amenable to rapid data prediction, yet allows many different classes of later arriving phases to be computed.

In some instances, *a priori* information is sufficiently detailed that more sophisticated parameterizations that mix continuous and discontinuous variations in seismic properties are warranted. For example, in exploration seismology, data coverage is usually dense, and near surface complexities (such as faults) often need to be accurately represented. Furthermore, there is often detailed information from field mapping and other geophysical techniques that is available. A parameterization that may be more suitable in these circumstances involves dividing the model region up into an aggregate of irregularly shaped volume elements (Fig. 4b), within which seismic properties vary smoothly, but is discontinuous across element boundaries (e.g. Pereyra, 1996; Bulant, 1999). This allows most geological features such as faults, folds, lenses, overthrusts, intrusions etc. to be faithfully represented, but makes both the forward prediction and inverse problems more challenging to solve.

The mathematical functions used to describe interfaces are largely analogous with those used to describe seismic continua. For example, piecewise linear segments are somewhat equivalent to constant velocity cells, and produce artificial ray shadow zones on account of the gradient discontinuities between each line segment (Williamson, 1990; Zelt and Smith, 1992). The logical extension of this to 3D is to represent surfaces using piecewise triangular area elements (Sambridge, 1990; Guiziou et al., 1996), as illustrated in

Fig. 4. Two schemes for representing media which contain both continuous and discontinuous variations in seismic property (a) laterally continuous interfaces within which seismic structure $w_i(x, z)$ varies smoothly; (b) flexible framework based on an aggregate of irregular blocks within which seismic structure $w_i(x, z)$ varies smoothly.
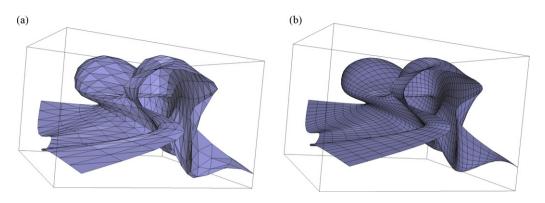
(a) (b)

Fig. 5. Multi-valued surface constructed using (a) a mesh of triangular area elements; (b) a mosaic of cubic B-spline surface patches.

Fig. 5a. Both of these parameterizations make multi-valued surfaces straightforward to represent, but cause two-point ray tracing and hence the data prediction problem to become less robust. Interfaces may also be defined on a grid of depth nodes, with some interpolation function used to describe the complete surface (Fig. 5b). Cubic splines are widely used in 2D and 3D wide-angle tomography, where sub-horizontal interfaces are commonly included (e.g. Farra and Madariaga, 1988; White, 1989; Lutter and Nowack, 1990; Pereyra, 1996; McCaughey and Singh, 1997; Rawlinson and Houseman, 1998; Rawlinson et al., 2001a; Rawlinson and Urvoy, 2006).
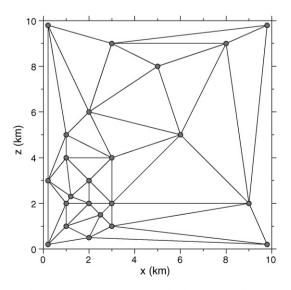
2.2. Irregular parameterizations

In regional and global tomography, it has long been recognised that the limited geographical distribution of sources and receivers often leads to highly irregular sampling of the subsurface by the recorded seismic energy. This problem also exists for more targeted experiments such as local earthquake and teleseismic tomography, although station distribution tends to be more uniform. Studies which control the location of sources, such as vibroseis, explosions, air-guns and ambient noise experiments, are less liable to experience uneven data coverage, but it still remains an issue. The use of uniform basis functions, as described above, to represent structural information extracted from such data is therefore inconsistent, because it does not recognise its spatially varying resolving power. An alternative approach is to use a parameterization which can itself adapt to the spatially varying constraints supplied by the data.

Pioneering work in this area goes back several decades, with the studies of Chou and Booker (1979) and Tarantola and Nercessian (1984), who propose "block-less" parameterizations for seismic tomography. These allow local smoothing scale lengths to vary spatially, and are in principle similar to the more recent and commonly used variable mesh schemes. "Continuous regionalization", as developed by Montagner and Nataf (1986), is one manifestation of the "block-less" approach to structural representation that is commonly used in surface wave tomography (e.g. Debayle, 1999; Debayle and Kennett, 2003). It produces a smooth model of variable scale length by using a Gaussian prior covariance function to enforce correlation between adjacent points. This takes the form of a prior variance and horizontal correlation length, which constrain the allowable amplitude and lateral length scale of anomalies. The choice of correlation length can be based on ray path coverage, which helps address the problem of uneven data sampling. One of the main drawbacks of the scheme is computational cost, which scales with M^2, where M is the number of data. Montagner and Tanimoto (1990) introduce several approximations to the original scheme to improve efficiency, and Debayle and Sambridge (2004) implement sophisticated geometrical algorithms to exclude

regions that contribute little to the prior covariance function. This has the dual benefit of further improving efficiency and making the algorithm highly suited to parallelization. Consequently, the new scheme can be applied to much larger problems (of the order of 50,000 paths for example).

In an alternative approach, Fukao et al. (1992) use non-uniformly sized rectangular 3D blocks to account for uneven ray sampling, and Abers and Roecker (1991) introduce a scheme in which fine scale uniform 3D blocks are joined to form larger irregular cells (a "bottom-up" approach). Sambridge et al. (1995) and Sambridge and Gudmundsson (1998) were the first to propose the use of Delaunay tetrahedra and Voronoi polyhedra, which are completely unstructured meshes, in seismic tomography (see Fig. 6 for an example of Delaunay triangulation—a continuum can be readily described for any arbitrary distribution of nodes). The main challenges in using such schemes include: (1) increased compute time to solve the forward problem; (2) developing an appropriate technique for fitting the mesh to the data constraints; (3) interpreting the results, which will exhibit structure at multiple scale lengths. Static schemes use a fixed parameterization throughout the inversion, while adaptive schemes dynamically adjust the parameterization during the inversion.

One of the first studies to use an adaptive scheme was that of Michelini (1995), who adjusts the velocity and position of cubic B-spline control vertices in 2D cross-hole tomography. While the topology of the control mesh in this case is regular, the use of parametric splines allows some irregularity in the position of

Fig. 6. Irregular parameterization using optimal Delaunay triangulation to describe a continuum based on a discrete set of control nodes.
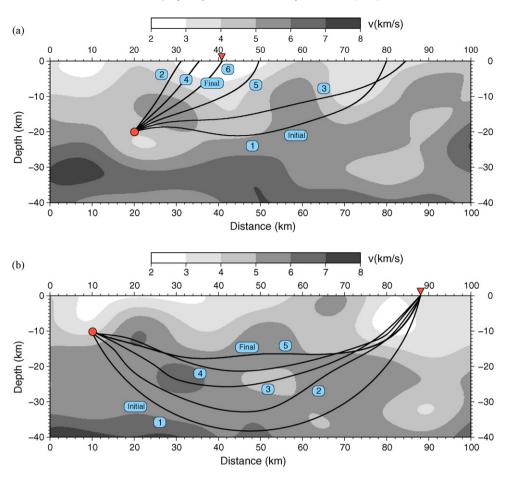
Fig. 7. Principle of the (a) shooting method; (b) bending method of ray tracing. In both cases, iterative refinement of some initial path is required to locate the correct two point path.

nodes. Curtis and Snieder (1997) also consider the 2D cross-hole problem, but use Delaunay triangulation to represent structure. A genetic algorithm is used to find the position of the node which minimizes the condition number of the tomographic system of equations. In 3D reflection tomography, Vesnaver et al. (2000) and Böhm et al. (2000) develop an adaptive scheme which uses Delaunay triangles and Voronoi polyhedra. Zhang and Thurber (2005) also devise an adaptive scheme based on tetrahedral and Voronoi diagrams to match the data distribution, and apply it to local earthquake and shot data to image the 3D structure beneath Parkfield, California.

Bijwaard et al. (1998), Bijwaard and Spakman (2000) and Spakman and Bijwaard (2001) perform global P-wave traveltime tomography using an approach similar to Abers and Roecker (1991) in which the 3D mesh is matched to the ray path density prior to inversion (i.e. a static approach). One of the first studies to carry out adaptive whole Earth tomography was that of Sambridge and Faletic (2003), who parameterize the Earth in terms of Delaunay tetrahedra. A "top-down" approach to mesh adaptation is used, in which new nodes are added to the edge of tetrahedra where the local velocity gradient is highest. Four updates are performed, with a linear tomographic system based on rays in a laterally homogeneous Earth solved after each update. This approach to adaptation is simple to implement, but regions of good data constraints are not always characterized by significant velocity gradients. Other studies to use Delaunay tetrahedra in global body wave tomography include those of Montelli et al. (2004) and Nolet and Montelli (2005). In fact, most global body wave imaging studies now use irregular meshes of one sort or another (e.g. Burdick et al., 2008).

A review of this topic can be found in Sambridge and Rawlinson (2005).

As noted earlier, spectral parameterizations such as spherical harmonics are not well suited to problems that exhibit significant variations in data coverage. An alternative approach that shows great promise in addressing the multi-scale nature of seismic tomography is the use of wavelet decomposition. Chiao and Kuo (2001) investigate the use of Harr wavelets on a sphere for representing lateral shear wave speed variations in D'', as constrained by S-SKS traveltimes. They conclude from their results that wavelets provide a natural regularization scheme based on ray path sampling, with recovered detail varying according to the resolving power of the data. Tikhotsky and Achauer (2008) invert both controlled source seismic and gravity data for 3D velocity and interface structure also represented using Haar wavelets. Loris et al. (2007) use more sophisticated wavelets that allow for smoother representations of structure than the discontinuous Haar wavelets. They also minimize an objective function that, in addition to the usual L_2 data misfit term, contains an L_1-norm measure of the wavelet coefficients, the aim being to promote a parsimonious description of structure that only has detail where required by the data.

A statistical method known as partition modelling, which is an ensemble inference approach used within a Bayesian framework, has recently been introduced to seismic tomography (Bodin and Sambridge, 2009). It uses a dynamic parameterization which is able to adapt to the uneven spatial distribution of information that characterizes most datasets, and does not require explicit regularization (damping and smoothing terms can be discarded). In the paper of Bodin and Sambridge (2009), a Markov chain Monte Carlo method

is used to invert traveltime data (assuming straight rays) with a model comprising a small number of constant velocity Voronoi cells. Remarkably, even though each model in the pool of best fitting solutions has a very blocky appearance, their average is a smooth model that is superior to that obtained using a conventional regular grid approach.

3. The data prediction problem

Seismic tomography may exploit one or more observables from a seismic record, including traveltime, amplitude, frequency content or the full waveform. The need for accurate, efficient and robust predictions of these quantities has driven the development of a wide range of techniques, most of which are based on the high-frequency assumption of geometric optics. The descriptions and discussion below will focus mainly on ray and grid-based techniques for solving the two-point problem of finding the path taken by seismic energy between source and receiver.

3.1. Ray-based methods

The full elastic wave equation can be greatly simplified in cases where the high frequency assumption is valid. It can be shown for both P and S waves in an isotropic medium (e.g. Rawlinson et al., 2007) that the elastic wave equation will reduce to:

$$|\nabla T| = s, \tag{1}$$

$$2\nabla A \cdot \nabla T + A\nabla^2 T = 0 \tag{2}$$

where T is traveltime, s is slowness and A is amplitude. Eq. (1) is the eikonal equation, which governs the propagation of seismic waves through isotropic media. Eq. (2) is the transport equation, which can be used to compute the amplitude of the propagating wave. In fully anisotropic media (Červený, 2001), the eikonal and transport equations have a slightly more complex form due to the presence of the elastic tensor **c**. Instead of directly solving the eikonal equation, it is possible to only consider its characteristics, which are trajectories orthogonal to the wavefront (in isotropic media). These are described by the kinematic ray equation:

$$\frac{d}{dl}\left[s\frac{dr}{dl}\right] = \nabla s. \tag{3}$$

where l is path length and **r** is a position vector of a point along the ray. In anisotropic media, the Hamiltonian formalism of classical mechanics (Červený, 2001; Chapman, 2004) is a more convenient form of representation. The behaviour of rays in the presence of interfaces is simply described by Snell's law, which can be generalized for anisotropic media (e.g. Slawinski et al., 2000). Dynamic ray tracing can be applied to yield amplitudes, and this can be done most easily by using the paraxial ray approximation (Červený and Pšenčík, 1983; Červený and Firbas, 1984; Červený, 1987; Farra and Madariaga, 1988; Virieux and Farra, 1991; Červený, 2001; Červený et al., 2007; Tian et al., 2007a), which essentially involves using first-order perturbation theory to deduce characteristics of the wavefield in the neighbourhood of a reference ray.

3.1.1. Shooting methods

Shooting methods of ray tracing formulate Eq. (3) as an initial value problem, which allows a complete ray path to be traced (with application of Snell's law at interfaces if necessary) given some initial trajectory. The two-point problem of locating a source–receiver path is more difficult to solve, because it is essentially a (potentially highly) non-linear inverse problem, with the initial ray direction as the unknown, and some measure of the distance between receiver and ray end point as the function to be minimized. In media described by constant velocity (or slowness) blocks, the initial value

problem is simple to solve (via repeated application of Snell's law), but the two-point problem is not (Williamson, 1990). Analytic ray tracing can also be used in media with a constant velocity gradient (e.g. White, 1989; Rawlinson et al., 2001a), constant gradient of ln v, and constant gradient of the nth power of slowness $1/v^n$ (Červený, 2001). Other than these few cases, numerical solution of Eq. (1) is required. In the presence of interfaces, one potentially difficult problem is to efficiently locate the ray-interface intersection point, particularly when sophisticated interface parameterizations are used. However, a number of practical methods are available (Sambridge, 1990; Virieux and Farra, 1991; Rawlinson et al., 2001a).

The boundary value problem is most commonly solved using an iterative non-linear approach, in which the source trajectory of some initial guess ray path is perturbed until it hits the desired end point (see Fig. 7a). Julian and Gubbins (1977) propose two different iterative non-linear techniques for solving the two-point problem: one is based on Newton's method, and the other on the method of false position. Both techniques have been widely used (e.g. Sambridge, 1990; Rawlinson et al., 2001a). Fig. 8 shows two point paths computed using the shooting method of Rawlinson et al. (2001a). A variety of methods have been proposed for locating a suitably accurate initial guess ray, including shooting a broad fan of rays towards the receiver and then iteratively refining the ray fan (Virieux and Farra, 1991), and using the correct two point path for a laterally averaged version of the model (Thurber and Ellsworth, 1980; Sambridge, 1990). As the non-linearity of the boundary value problem increases, iterative non-linear solvers require more accurate initial guess rays (see Fig. 10 of Rawlinson et al., 2007, for a clear illustration). Although not frequently acknowledged in the literature, practical applications of shooting, particularly in regions of significant heterogeneity, often settle for some "acceptable" trade-off between the percentage of two-point paths located, and total compute time.

Fully non-linear shooting methods, based on sampling algorithms like simulated annealing, have been devised and tested (e.g. Velis and Ulrych, 1996, 2001), but they have not proved popular. Perhaps this is because ray tracing is at its most useful when velocity heterogeneity is not too severe, so that local sampling of the ray field is still a valid approach for the detection of two-point paths. When this is no longer the case, global techniques like grid based eikonal solvers (see below) will be much more efficient. Shooting methods of ray tracing are widely used in seismic tomography, due to their conceptual simplicity, and potential for high accuracy and efficiency (Cassell, 1982; Benz and Smith, 1984; Langan et al., 1985; Farra and Madariaga, 1988; Sambridge, 1990; Zelt and Smith, 1992; VanDecar et al., 1995; McCaughey and Singh, 1997; Rawlinson et al., 2001b).

3.2. Bending methods

Bending methods of ray tracing iteratively adjust the geometry of some arbitrary two point path until it becomes a true ray path (see Fig. 7b) i.e. it satisfies Fermat's principle of stationary time. A common approach to implementing the bending method is to derive a boundary value formulation of Eq. (3), which can then be solved iteratively (Julian and Gubbins, 1977). Pereyra et al. (1980) devise a bending method similar to that of Julian and Gubbins (1977), but extend it to allow for the presence of interfaces. In complex media Pereyra (1996) use ray shooting to help locate an initial guess ray.

Pseudo-bending methods use the same principle of adjusting ray geometry to locate a true ray, but avoid direct solution of the ray equations. One of the first pseudo-bending schemes was developed by Um and Thurber (1987), who describe a ray path by a set of linearly interpolated points. For some initial arbitrary path described by a small number of points, the scheme proceeds by
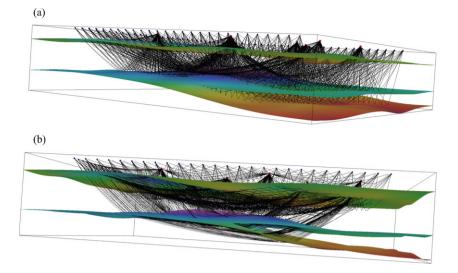
Fig. 8. Two point paths through a 3D medium computed using the shooing method of Rawlinson et al. (2001a). (a) Reflected rays; (b) refracted rays.

adjusting the location of each point by directly exploiting Fermat's principle of stationary time. Once some convergence criterion is satisfied, new points are interpolated between pre-existing points, and the iterative procedure continues until sufficient accuracy is achieved. Despite the relatively crude approximations made in pseudo-bending, Um and Thurber (1987) find it to be much more computationally efficient than conventional bending schemes. Consequently, it has become popular for problems that require large travel-time datasets to be predicted, such as in 3D local earthquake tomography (Eberhart-Phillips, 1990; Scott et al., 1994; Eberhart-Phillips and Reyners, 1997; Graeber and Asch, 1999). Zhao et al. (1992) modify the pseudo-bending scheme of Um and Thurber (1987) to allow for the presence of interfaces, and Koketsu and Sekine (1998) devise a similar scheme in 3D spherical coordinates.

Like ray shooting, fully non-linear bending methods have also been devised; for example, Sadeghi et al. (1999) develop a method which uses genetic algorithms to globally search for the minimum time path between two fixed points. Again, like fully non-linear shooting, one could argue that the exhaustive interrogation of the ray field for each source–receiver pair would make other classes of techniques that guarantee to find the global minimum (like eikonal solvers) more practical. Apart from shooting and bending methods, the boundary value problem can also be solved using techniques based on structural perturbation (Červený, 2001). In this class of scheme, a known two-point path exists in a reference medium, and the aim is to locate the equivalent two point path in a slightly modified medium. Solution of this class of problem can be obtained using ray perturbation theory (Farra and Madariaga, 1987; Snieder and Sambridge, 1992; Snieder, 1993; Pulliam and Snieder, 1996). Although relevant to iterative non-linear tomography, this approach is not widely used.
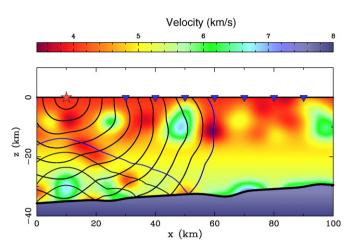
3.3. Grid-based methods

Instead of tracing rays between source and receiver, an alternative strategy is to compute the global traveltime field as defined by a grid of points. This will implicitly contain the wavefront geometry as a function of time (i.e. contours of $T(\mathbf{x})$), and all possible ray trajectories (specified by ΔT). Compared to conventional shooting and bending methods of ray tracing, grid-based traveltime schemes have several clear advantages: (1) they compute traveltimes to all points in the medium, including (in most cases) diffractions in ray shadow zones; (2) they exhibit high stability in strongly heterogeneous media; (3) they efficiently compute traveltime and path

information, particularly when the ratio of sources to receivers (or vice versa) is high; (4) they consistently yield first-arrivals. The advantages of grid-based schemes are offset somewhat by the following drawbacks: (1) their accuracy is a function of grid-spacing—in 3D, halving the grid spacing will increase compute time by at least a factor of eight; (2) in most cases, they only produce first-arrivals; (3) they have difficulty computing quantities other than traveltime (e.g. amplitude); (4) anisotropy, easily dealt with by ray methods, is more of a challenge. Two grid-based schemes – finite difference solution of the eikonal equation and shortest path methods – have emerged as popular alternatives to ray tracing.

3.3.1. Eikonal solvers

The use of eikonal solvers in seismology was largely pioneered by Vidale (1988, 1990), who developed a technique for finite difference solution of the eikonal equation on a grid. Relatively simple centred difference stencils are formulated for approximating the gradient terms in Eq. (1), so that traveltimes can be computed at new points using known values at adjacent points. An expanding square is adopted for the computational front, which sweeps through the medium from the source point until the complete traveltime field is computed. Ray paths can be found retrospectively by simply following the traveltime gradient from each receiver back to the source. The resulting scheme is fast and sufficiently accurate for most seismic applications, with CPU time being approximately proportional to the number of points defining the grid. The use of an expanding square formalism to define the shape of the computational front cannot always respect the direction of flow of information through the medium. For example, it is possible that a first arrival will need to sample a region outside the expanding square before returning back into it. Consequently, first arrivals are not always guaranteed, which can lead to instability. Nonetheless, the basic scheme proposed by Vidale (1988, 1990) remains popular, and its stability has been improved thanks to new features such as special headwave operators (Hole and Zelt, 1995; Afnimar and Koketsu, 2000), and post-sweeping to correct for the non-causal nature of the expanding square (Hole and Zelt, 1995). van Trier and Symes (1991) use entropy-satisfying first-order upwind operators to improve computational efficiency and deal with wavefront discontinuities. Comparable improvements are made by Podvin and Lecomte (1991), who employ Huygen's principle in the finite difference approximation.

The above techniques, which have largely been independently developed in seismology, bears some similarity with essentially

Fig. 9. Reflected wavefront and traveltime field computed using the FMM scheme of Rawlinson and Sambridge (2004a).

non-oscillatory (ENO) finite difference schemes developed in the field of computational mathematics (Shu and Osher, 1988, 1989). The attraction of ENO schemes is that they can be readily expanded to very high orders of accuracy, yet remain stable. Kim and Cook (1999) devise a scheme which they describe as ENO–DNO–PS to efficiently compute first-arrival traveltime fields. The DNO (or down 'n' out) refers to an expanding box computational front, and PS refers to post-sweeping. Therefore, apart from the finite difference stencil, it is very similar to the scheme of Hole and Zelt (1995). WENO (weighted ENO) schemes, which offer improved computation time and stability compared with ENO schemes, have also been developed (Liu et al., 1994; Jiang and Shu, 1996; Jiang and Peng, 2000). Qian and Symes (2002) use a WENO scheme with adaptive gridding to compute traveltimes, and Buske and Kästner (2004) implement a WENO scheme in polar coordinates to compute traveltimes that are sufficiently accurate to solve the transport equation for amplitudes.

In order to overcome the limitations of the expanding square formalism, Qin et al. (1992) and Cao and Greenhalgh (1994) use the first-arriving wave-front as the computational front, which is locally evolved by always choosing the point with minimum traveltime along the edge of the computed field to update adjacent points. This ensures that the shape of the computational front conforms to the first-arrival wavefront, and will not result in causality breaches. The drawback of this approach is the additional computational expense involved in locating the global minimum traveltime point along the computational front. For example, if heap sorting is used, then computing time will be proportional to $N \log N$, where N is the total number of grid points used to describe the velocity field.

Another eikonal solver that was developed in the field of computational mathematics is the so-called Fast Marching Method or FMM (Sethian, 1996; Sethian and Popovici, 1999; Popovici and Sethian, 2002). It uses upwind entropy-satisfying finite difference stencils to solve the eikonal equation, and a computational front (narrow band) that encapsulates the first-arriving wavefront. The finite difference stencils account for the fact that the first-arrival traveltime field is not always differentiable (i.e. the ΔT term in Eq. (1) is not necessarily defined everywhere), and result in a very robust method. Rawlinson and Sambridge (2004a,b) extend the scheme to improve accuracy in the source neighbourhood where wavefront curvature is high (and therefore poorly described by a regular grid), and compute phases comprising any number of refracted and reflected branches in media containing interfaces (see Fig. 9). Eikonal solvers are now widely used in tomography, particularly 3D wide-angle and teleseismic studies (Hole, 1992; Zelt et

al., 1996, 2001; Riahi et al., 1997; Zelt and Barton, 1998; Zelt, 1999; Day et al., 2001; Rawlinson et al., 2006a,b; Rawlinson and Urvoy, 2006; Rawlinson and Kennett, 2008).

3.3.2. Shortest path ray tracing

Another class of grid-based scheme that has been used in seismic tomography to compute traveltimes to all points of a gridded velocity field is shortest path ray tracing or SPR (Nakanishi and Yamaguchi, 1986; Moser, 1991; Fischer and Lees, 1993; Cheng and House, 1996; Zhao et al., 2004; Zhou and Greenhalgh, 2005). Instead of solving a differential equation, a network or graph is formed by connecting neighbouring nodes with traveltime path segments. Dijkstra-like algorithms can then be used to find the shortest time path between a source and receiver, which, according to Fermat's principle of stationary time, will correspond to a valid ray path. Shortest path networks are commonly defined in terms of either a cell or grid centred framework. In the latter case, the connection stencil is often referred to as the "forward star" (Klimeš and Kvasnička, 1994). In 2D a forward star with 8 connections will join any node with all of its immediate neighbours, but will not allow variations in ray path trajectory less than 45° (for a square grid). By allowing direct connections between the central node and the neighbours of the 8 nodes, a forward star with 16 connections can be defined, which will permit greater flexibility of the ray geometry.

The difference between a scheme like FMM and SPR is actually not all that great; they both use the shape of the first-arriving wavefront as the computational front and use the same approach for choosing a new node for a local update of the traveltime field. The only obvious change is in the update stencil that is used to compute new traveltimes. As such, many of the extensions that have been applied to FMM, such as grid refinement and the location of later arriving phases consisting of reflected and refracted branches, are equally applicable to SPR (Moser, 1991). Although less frequently used than eikonal solvers, SPR has been implemented in a number of tomographic studies to solve the forward problem of traveltime prediction (Nakanishi and Yamaguchi, 1986; Toomey et al., 1994; Zhang and Toksöz, 1998; Bai, 2005).

3.4. Multi-arrival schemes

All of the schemes described above are really only suitable for tracking a single (usually the first) or a limited number of arrivals between two points. However, even relatively small variations in seismic wavespeed can produce a phenomenon known as multi-pathing, which is simply defined as when more than one ray path connects two points in the medium. In order for this to occur, the propagating seismic wavefront must distort to such an extent that it self-intersects (or folds over on itself). In such situations, eikonal and shortest path methods will yield the first arrival only, while most shooting and bending methods will only produce a single arrival (with no real guarantee as to whether it is a first or later arrival). With sufficient probing of the medium, multiple arrivals can be produced with ray tracing, but usually not in a robust or efficient enough manner for applications such as tomography. The heterogeneity of the Earth, particularly near the surface, means that multi-pathing commonly contributes to the complexity of recorded waveforms. Therefore, any method that can accurately predict all arrivals of significant amplitude has important implications for Earth imaging.

To date, a number of methods have been developed to solve the multi-arrival problem. These include both grid-based (Benamou, 1999; Steinhoff et al., 2000; Engquist et al., 2002; Fomel and Sethian, 2002; Osher et al., 2002; Symes and Qian, 2003), ray based (Vinje et al., 1993; Lambaré et al., 1996; Vinje et al., 1996, 1999; Hauser et al., 2008) and hybrid (Benamou, 1996) schemes. Due to the relative infancy of grid-based schemes, which have been largely

developed outside the field of seismology, most of the methods cited above are not sufficiently developed for practical application in tomography. However, recent work by Cheng (2007) which develops and refines the level-set approach of Osher et al. (2002) indicates that it is now on the point of being computationally feasible for seismic tomography.

In contrast to grid-based schemes, ray-based schemes for tracking multi-arrivals are much more mature, and have been used in seismology. One of the first ray-based schemes was proposed by Vinje et al. (1993), and is commonly referred to as "wavefront construction". The basic principle underlying this technique is that a wavefront can be discretely described by a set of points, which are then progressively evolved through a medium using initial value ray tracing for a given time increment. In order to avoid undersampling of the wavefront as it focuses and defocuses, new points are inserted based on a distance criterion between adjacent rays. The original 2D method of Vinje et al. (1993) was subsequently extended to 3D (Vinje et al., 1996), and then modified for media involving complex interfaces (Vinje et al., 1999). Since the initial value ray tracing equations can be solved with high accuracy, the main source of error comes from the interpolation step. Simply using a distance metric in normal space can introduce instability in this process, because wavefronts may contain significant kinks, particularly when they triplicate. Lambaré et al. (1996) and Hauser et al. (2008) use an interpolation criterion based on the phase space distance between adjacent points to overcome this problem. Wavefront construction has been used in various applications including coincident reflection migration (Xu and Lambaré, 2004; Xu et al., 2004). Hauser et al. (2008) investigate its use in the context of multi-arrival seismic tomography, and conclude that it has the potential to significantly improve the quality of recovered images, particularly in regions of low wavespeed which are poorly constrained by first-arrival ray paths.

3.5. Finite frequency considerations

Geometric ray theory has been an integral part of seismic tomography for the past four decades, and can be used to predict various quantities including traveltime, geometric spreading amplitude and t^* (used in attenuation tomography). However, seismic tomography that accounts for finite frequency effects is starting to become more common, due to the increasing recognition of its importance for accurate imaging, and continuing improvements in computing power. This is on top of the gradual emergence of techniques that attempt to invert the full seismic waveform, and implicitly account for finite frequency effects by numerically solving the wave equation.

In the last decade or so, the term "finite frequency tomography" has come to mean tomography that employs first-order perturbation theory (or Born scattering theory) to account for scattering/diffraction effects including wavefront healing. Most implementations of this technique exploit the frequency dependence of traveltime that arises largely from the wavefront healing phenomenon (Nolet, 2008) and build the forward calculation on top of kinematic and dynamic ray tracing. Dahlen et al. (2000) comprehensively describes the theory for body waves in 3D media. The focus of this paper is the efficient calculation of Fréchet kernels (or sensitivity kernels), traditionally one of the impediments to successful finite frequency body wave tomography. In order to fully account for the first-order dependence of absolute or differential traveltimes (measured using waveform cross-correlation) on structural perturbations, not only does the source–receiver geometric ray need to be found, but so do all possible rays from the source to every scatterer, and all possible rays from the receiver to every scatterer. Furthermore, traveltimes, Maslov indices, reflection-transmission products, geometrical spreading factors and other

quantities need to be computed along each of these rays. The computational burden of this approach makes it impractical for routine applications of seismic tomography.

Dahlen et al. (2000) develop an alternative scheme which only requires a single geometric ray path to be computed for each source–receiver pair. This approximation is generally valid because perturbations in traveltime tend to be only sensitive to structural perturbations in the immediate neighbourhood of the central geometric ray i.e. approximately the first Fresnel zone. Ignoring scatterers outside this region (and along the geometric ray, which according to first-order scattering theory has no influence on cross-correlated traveltime measurements), vastly improves computation time, and enables paraxial ray theory to be invoked to increase efficiency even further. Using the paraxial approximation only requires the geometric ray to be traced because information in the neighbourhood of the central ray can be readily computed. Thus, in addition to finding the single two-point ray, only a few additional integrations along the ray path are necessary to build the sensitivity kernel. Dahlen et al. (2000) extend their new theory to cover overlapping phases such as direct P and pP that arrive at similar times. In a companion paper, Hung et al. (2000) illustrate several examples of both absolute and differential traveltime sensitivity kernels and Montelli et al. (2004) apply the new theory to global body wave mantle tomography. Calvet and Chevrot (2005) develop an alternative scheme for computing PKP kernels that rely on look-up tables for traveltimes and geometric spreading. Although more time consuming, this approach is preferable to a paraxial approach when kernels become very wide (Tian et al., 2007b). Finite frequency sensitivity kernels have also been derived for surface waves (e.g. Marquering et al., 1998; Zhou et al., 2004).

One of the common assumptions made in finite frequency tomography is that scattering from one wave type to another (e.g. S to P) is not significant and can be ignored. However, in a recent study, Zhang and Shen (2008) use finite difference solution of the wave equation to demonstrate that S-wave perturbations can have a significant effect on P waveforms, and conclude that ignoring this cross-dependence can lead to systematic bias in the recovery of P wavespeeds.

Exploiting the full waveform in seismic tomography requires an efficient method for solving the elastic wave equation. The problem can be formulated and solved in a variety of ways, all of which are, at least compared to geometric ray methods, computationally intensive. The far field elastic wave equation for isotropic media can be written as:

$$\rho\ddot{\mathbf{u}} = \nabla\lambda(\nabla \cdot \mathbf{u}) + \nabla\mu \cdot [\nabla\mathbf{u} + (\nabla\mathbf{u})^{\mathsf{T}}]$$
$$+ (\lambda + 2\mu)\nabla(\nabla \cdot \mathbf{u}) - \mu\nabla \times \nabla \times \mathbf{u} \qquad (4)$$

where **u** is displacement, $\ddot{\mathbf{u}}$ is the second derivative of displacement with respect to time (i.e. acceleration), ρ is density, and λ and μ are the Lamé parameters. Eq. (4) can be expressed in a variety of ways depending on the nature of the source, the frequency content of the seismic waves and the assumptions made about the medium. As an initial value partial differential equation, Eq. (4) can be solved using finite difference or finite element methods.

The potential for solving the wave equation in the context of full waveform tomography has long been recognised, particularly in active source studies such as cross-hole and reflection imaging. Early efforts focused on the acoustic problem (Tarantola and Nercessian, 1984; Pratt and Worthington, 1990), which is simpler to deal with but has limited application due to source generated and mode-converted shear waves. The advantages of solving the wave equation in the frequency domain were also understood early on, with the development of both finite element and finite difference schemes (Marfurt, 1984; Pratt, 1990). Compared to time domain schemes (e.g. Virieux, 1984, 1986), frequency domain methods

can solve the wave equation for additional source positions with minimal added cost, and are more efficient when only a limited number of frequency components are required, which in practice is often the case in seismic tomography. Pratt (1990) reformulates Eq. (4) in the frequency domain for 2D media, and incorporates attenuation by making the Lamé parameters complex valued and frequency dependent. The resultant system of elliptic partial differential equations constitutes a boundary value problem, which is solved using absorbing boundary conditions to eliminate artificial reflections. The use of 2D models result in a more tractable forward problem, but the dissipation of elastic energy from the source in the out-of-plane direction is difficult to calculate. Solving the forward problem in "2.5D" media (i.e. models that are 3D, but only have structural variations in two dimensions) can mitigate this shortcoming without the computational expense of a fully 3D wave equation solver (e.g. Song et al., 1995). Since this early work, increasing computer power and the development of more efficient wave equation solvers (e.g. Štekl and Pratt, 1998) have given rise to more ambitious applications of full waveform tomography. For example, Sirgue and Pratt (2004) incorporate frequency domain solution of the elastic wave equation in their full waveform inversion of seismic refraction data to recover the detailed structure of the complex Marmousi model.

The development and application of full waveform tomography in passive source studies has been a more recent phenomenon, a fact which can probably be attributed to the large data volumes that are often involved, the 3D nature of many of these experiments, and the more complex and less well understood nature of the source mechanism. Surface waveform tomography has been around for a long time (e.g. Nolet, 1990), but generally only use 1D waveform inversion of long period waves, which are then combined to form a 3D model. Full waveform simulation methods for local, regional and global models have been in existence for some time, but tend to be computationally expensive, and have difficulties dealing with free-surface boundary conditions (e.g. Frankel and Vidale, 1992; Carcione, 1994; Graves, 1996; Faccioli et al., 1997; Furumura et al., 1998). Finite difference techniques are conceptually straightforward to implement, but with at least 5–7 grid points per minimum wavelength required for sufficient accuracy, extremely large grids are required to propagate anything other than very low frequency waves. Finite difference techniques, which generally require a regular grid, are also not well suited to an irregular free surface. Finite element techniques can overcome this limitation, but are still computationally expensive. An alternative strategy that uses spectral approximations of the displacement field is developed by Faccioli et al. (1997); advantages over the more traditional techniques include fewer grid points per wavelength, accommodation of complex geometries through the use of irregular hexahedral volume elements, and natural partitioning into subdomains that favour a parallel implementation.

In the context of developing a practical 3D passive source full waveform tomography technique, the current wave equation solver of choice appears to be the spectral element method or SEM (Komatitsch and Vilotte, 1998; Komatitsch and Tromp, 1999; Komatitsch et al., 2002; Liu and Tromp, 2008), which was originally developed in the field of computational fluid dynamics. Like the pseudo-spectral method of Faccioli et al. (1997), it attempts to combine the flexibility of finite element methods with the accuracy of spectral methods. In fact, the differences between the two techniques do not appear to be very great; for instance, hexahedral elements, which can adapt to an irregular free surface, are used by both methods, as is a similar Legendre Gauss–Lobatto quadrature approach to numerical integration within each volume element. However, SEM has been developed to the point that it can be applied at a variety of scales from local to global, and can account for a range of physical phenomena including anelasticity, anisotropy,

rotation of the Earth, self-gravitation, presence of the oceans, etc. (Komatitsch et al., 2005). Software for solving the full wave equation, and computing kernels for finite frequency tomography, can be found at the CIG (Computation Infrastructure for Geo-dynamics) website: http://geodynamics.org/cig/software/packages/seismo.

4. Solving the inverse problem

The inversion step in seismic tomography requires the adjustment of model parameters $\mathbf{m}$ to satisfy data observations $\mathbf{d}_{obs}$, subject to any independent constraints (commonly referred to as regularization). One of the main difficulties to be addressed in this procedure is that of solution non-uniqueness, which pervades all practical applications of seismic tomography. Most published studies end up interpreting a single model, but how can this be justified given that a range of models invariably satisfies the data to the same extent? The appeal of regularization is that it considerably reduces the size of the subspace that contains data-fitting solutions, making subsequent selection of a preferred model more straightforward. The drawback is that *ad hoc* constraints are often imposed. Various techniques also exist for analyzing the robustness of solution models or model ensembles. Many, such as the ubiquitous checkerboard test discussed later, lack rigour, but given the significant dimension of most problems, more comprehensive methods are often not computationally feasible.

In addition to solution non-uniqueness, non-linearity of the inverse problem also afflicts many seismic tomography applications. This is certainly true of the most common form of tomography, which uses source–receiver traveltimes to build images of velocity heterogeneity. Traveltime and slowness (inverse of velocity) are linearly related for a given path, but the path itself has a non-linear dependence on the velocity field, which changes as a result of the inversion. Thus:

$$t = \int_{L(\upsilon)} \frac{1}{\upsilon(\mathbf{x})} \mathrm{d}l \tag{5}$$

where L is the ray path and $\upsilon(\mathbf{x})$ is the velocity field. Although not strictly based on the above equation, both finite frequency and full waveform tomography are also non-linear (clearly, the relationship between displacement and the Lamé parameters in Eq. (4) is non-linear). In some cases, the inverse problem can be treated as linear. For example, attenuation tomography often uses the relationship between the attenuation parameter t^* and the quality factor Q as (e.g. Pozgay et al., 2009):

$$t^* = \int_{L(\upsilon)} \frac{1}{\upsilon(\mathbf{x})Q(\mathbf{x})} \mathrm{d}l \tag{6}$$

which ignores any ray perturbation effects due to attenuation. Generally, attenuation tomography is performed subsequent to traveltime tomography, which means that for a given velocity solution model, the relationship between t^* and Q^{-1} is linear.

One interesting question that is not often discussed is what constitutes a data-satisfying solution model in seismic tomography? In most published studies, it is likely that at least some subset of data predictions do not match their observational counterparts within error estimates. A common statistical test that is sometimes used is the so-called χ^2 test, which is simply the sum of the square of differences between each observation and prediction weighted by the corresponding data uncertainty. When the normalized χ^2 value drops below one, then one could regard the model as satisfying the data (although some observations may still be poorly matched). The potential usefulness of this test is tempered by the fact that data uncertainty is often poorly known; the implicit regularization imposed by the assumed parameterization, coupled with the need to regularize the inversion, means that it may no longer be possi-
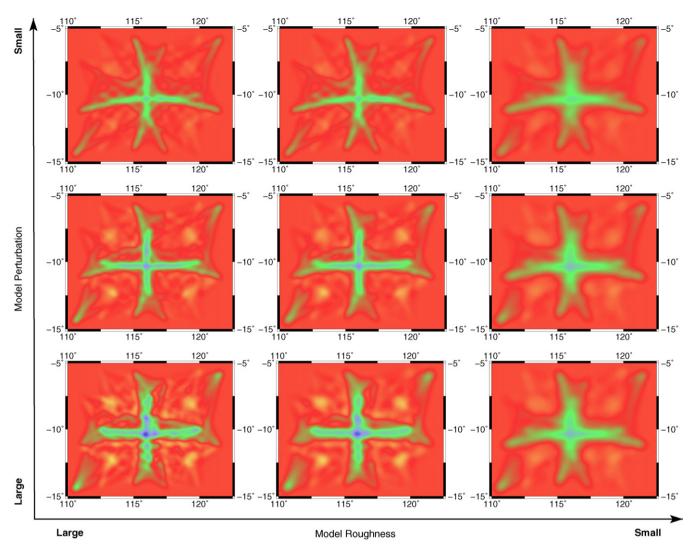
Fig. 10. Illustration of the effects of smoothing and damping on the solution model shown in Fig. 1. The model roughness is varied using the smoothing factor η, and the model perturbation is controlled using the damping parameter ε.

ble to achieve a χ^2 value of one; and the use of approximations in the theory (e.g. geometric ray theory) will introduce error into the predictions. Thus, the "solution model" that is often produced in published papers is one that achieves some balance between data fit (according to some measure such as χ^2 or RMS misfit) and the influence of regularization (e.g. smoothness, deviation from initial model), without strictly "satisfying" the data.

4.1. Backprojection

Backprojection methods exploit the fact that most data observations (e.g. traveltime) can be viewed as integral quantities along a ray path. The basic aim is to project this measured quantity back along the ray path from the receiver to the source and use the known relationship $\mathbf{d} = \mathbf{g}(\mathbf{m})$ to convert incremental data values to local model perturbations. In traveltime tomography, the technique is generally used in conjunction with constant slowness blocks. Thus, the relationship between traveltime perturbations $\mathbf{d}$ and slowness perturbations $\mathbf{m}$ can be linearized as $\mathbf{d} = \mathbf{Gm}$ where $\mathbf{G}$ is a matrix of ray lengths corresponding to the distance traversed by each ray in each block. An initial model that is relatively close to the solution model is required to justify the assumption of local linearity. Many of the elements of $\mathbf{G}$ will be zero since each ray path will usually only traverse a small subset of blocks. The basic aim of

backprojection is to solve the linear system of equations described by $\mathbf{d} = \mathbf{Gm}$, but not directly.

Two well-known backprojection methods are the algebraic reconstruction technique (ART), and the simultaneous iterative reconstruction technique (SIRT), both of which originate from medical imaging. ART is the more basic of the two techniques, because it simply updates the model on a ray-by-ray basis. Each residual is distributed along its associated path by adjusting the slowness in each cell in proportion to the length of the ray path segment in each cell. Before repeating this process for the next ray, the traveltime residual is computed from the updated velocity field. Once the procedure has been carried out for all rays, the inversion is either complete (linear tomography), or new ray paths can be found for the updated model, and the backprojection repeated. This cycle can be repeated until the traveltime residuals satisfy some convergence criterion (iterative non-linear tomography). The main problem with ART is that it suffers from poor convergence properties (Blundell, 1993), but it has found application in cross-hole and local earthquake tomography (e.g. McMechan, 1983; Nakanishi and Yamaguchi, 1986).

Rather than sequentially update the model on a ray-by-ray basis, SIRT averages the perturbations applied to each cell from all traversing rays. This more sophisticated approach yields an algorithm with superior convergence properties, which perhaps

explains its more frequent use in seismic tomography (e.g. Granet and Trampert, 1989; Dueker et al., 1993; McQueen and Lambeck, 1996). Variants of these backprojection schemes that include some form of regularization have also been developed and applied in seismic tomography (e.g. Humphreys and Clayton, 1990; Hole, 1992; Zelt and Barton, 1998). Humphreys and Clayton (1990) apply filtering and spatial averaging to reduce blurring and produce a smooth solution, and Zelt and Barton (1998) implement several other modifications aimed at improving the convergence and accuracy of backprojection. The attractiveness of backprojection is that it is simple to implement and computationally rapid at each iteration. However, it can suffer from stability and convergence problems, and does not naturally lend itself to regularization. Probably for these reasons, backprojection is now infrequently encountered in published seismic tomography studies.

4.2. Gradient methods

Inversion methods that use the derivative of model predictions ($\partial \mathbf{g}/\partial \mathbf{m}$) in order to produce a solution are by far the most common in seismic tomography. These schemes are often applied within a formal framework that involves the minimization of an objective function containing a data residual term and one or more regularization terms. For example, a typical objective function might look like (Rawlinson et al., 2006b):

$$S(\mathbf{m}) = (\mathbf{g}(\mathbf{m}) - \mathbf{d}_{\text{obs}})^{\text{T}} \mathbf{C}_d^{-1} (\mathbf{g}(\mathbf{m}) - \mathbf{d}_{\text{obs}})$$
$$+ \varepsilon (\mathbf{m} - \mathbf{m}_0)^{\text{T}} \mathbf{C}_m^{-1} (\mathbf{m} - \mathbf{m}_0) + \eta \mathbf{m}^{\text{T}} \mathbf{D}^{\text{T}} \mathbf{D} \mathbf{m} \quad (7)$$

where $\mathbf{g}(\mathbf{m})$ are the predicted residuals, $\mathbf{d}_{\text{obs}}$ are the observed residuals, $\mathbf{C}_d$ is the *a priori* data covariance matrix, m_0 is the reference model, $\mathbf{C}_m$ is the *a priori* model covariance matrix, and $\mathbf{D}$ is a second derivative smoothing operator. ε and η are known as the damping parameter and smoothing parameter respectively, and govern the trade-off between how well the solution satisfies the data, the proximity of the solution model to the starting model, and the smoothness of the solution model. The last two terms on the right hand side of Eq. (7) are regularization terms, which have the effect of limiting the number of acceptable data-fitting models.

Although commonly used, Eq. (7) effectively juxtaposes two different regularization frameworks: Bayesian and Occam's. In a Bayesian style inversion, information is represented in probabilistic terms; prior information on a model is combined with constraints supplied by the data to produce a posterior distribution. For a strict implementation of this approach, η would be set to zero in Eq. (7), and $\mathbf{C}_d$ and $\mathbf{C}_m$ would accurately reflect the uncertainty associated with the data and initial model respectively. It is also necessary to set $\varepsilon = 1$, since it is not meaningful to re-weight the prior uncertainty. Minimization of the objective function has the effect of assimilating the prior information with the data constraints, producing a model with posterior uncertainties that are smaller than the prior uncertainties (how much depends on how good the data are). The main impediment to the success of this approach is that meaningful information on prior model and data uncertainties is difficult to obtain in practice. An example of the difficulties of choosing an appropriate prior model is illustrated in Fishwick et al. (2005). In surface wave studies of the upper mantle a global reference model (e.g. PREM or ak135) is often chosen as the prior model; however at depths of 100–200 km it is likely that these models are not very representative of either oceanic or cratonic regions, and damping back towards the reference model is likely to underestimate the true variations in wavespeed. Scales and Snieder (1997) discuss the merits and pitfalls of adopting a Bayesian framework for inversion.

The other regularization framework is suggested by Occam's Razor, which favours parsimony over complexity when arriving at

a hypothesis to fit the observations. In seismic tomography, this will manifest as the solution with the least structure necessary to fit the data (Constable et al., 1987). In this case, one would set $\varepsilon = 0$, because damping back to the initial model may exclude the minimum data-satisfying structural model. This type of scheme is commonly used in seismic tomography (Sambridge, 1990; Zelt and Barton, 1998; Day et al., 2001), although it should be noted that the second spatial derivative, as in Eq. (7), is only one measure of model complexity. For instance, in the frequency domain, one could seek the solution with the least number of harmonic terms that satisfies the data.

In most cases, a mixture of both frameworks is used, in which the aim is to find a physically reasonable model that contains no unnecessary structure, is in the neighbourhood of the initial model, and satisfies the data. When this occurs, $\mathbf{C}_d$ and $\mathbf{C}_m$ are not strictly required to be covariance matrices, since the presence of the free parameters ε and η make their absolute values meaningless. Fig. 10 demonstrates the effect of damping and smoothing on the solution model shown in Fig. 1. When minimal smoothing and damping are used, the recovered structure contains numerous short wavelength artifacts. Damping tends to decrease the amplitude of perturbations without filtering the image, while smoothing essentially acts as a low pass filter. In addition to the explicit smoothing imposed by Eq. (7), implicit smoothing via the use of a cubic B-spline parameterization is also present. There are various ways for choosing the "optimum" damping and smoothing parameters, the most common of which is to plot trade-off curves between data fit and model perturbation and roughness (Rawlinson and Sambridge, 2003a). Ideally, these plots will have an "L" shape, with the elbow of the curve indicating the optimum trade-off region. More statistically rigorous methods such as generalized cross-validation can also be used to determine ε and η (e.g. Lukas, 2008).

4.2.1. Solution strategies

Gradient based inversion methods make use of the derivatives of the objective function at a specified point in model space under the assumption that $S(\mathbf{m})$ is sufficiently smooth to permit a local quadratic approximation:

$$S(\mathbf{m} + \delta\mathbf{m}) \approx S(\mathbf{m}) + \gamma\delta\mathbf{m} + \tfrac{1}{2}\delta\mathbf{m}^{\text{T}}\mathbf{H}\delta\mathbf{m} \quad (8)$$

where $\delta\mathbf{m}$ is a perturbation to the current model and $\gamma = \partial S/\partial \mathbf{m}$ and $\mathbf{H} = \partial^2 S/\partial \mathbf{m}^2$ are the gradient vector and Hessian matrix respectively. The goal of gradient-based inversion methods is to determine $\delta\mathbf{m}$; in cases where the function $\mathbf{g}$ is non-linear, minimization of Eq. (7) requires an iterative approach:

$$\mathbf{m}_{n+1} = \mathbf{m}_n + \delta\mathbf{m}_n \quad (9)$$

where $\mathbf{m}_0$ is the initial model. The forward data prediction problem (e.g. recomputing of ray paths) is solved after each model update, and the process concludes when the data are satisfied or some convergence criterion is met.

There are many different methods available for computing the model perturbation $\delta\mathbf{m}$. The classic Gauss–Newton method computes the model update by finding the minimum of the tangent paraboloid to $S(\mathbf{m})$ at $\mathbf{m}_n$, which produces:

$$\delta\mathbf{m}_n = -[\mathbf{G}_n^T \mathbf{C}_d^{-1} \mathbf{G}_n + \nabla_m \mathbf{G}_n^T \mathbf{C}_d^{-1} (\mathbf{g}(\mathbf{m}_n) - \mathbf{d}_{\text{obs}}) + \varepsilon \mathbf{C}_m^{-1} + \eta \mathbf{D}^T \mathbf{D}]^{-1}$$
$$\times [\mathbf{G}_n^T \mathbf{C}_d^{-1} [\mathbf{g}(\mathbf{m}_n) - \mathbf{d}_{\text{obs}}] + \varepsilon \mathbf{C}_m^{-1} (\mathbf{m}_n - \mathbf{m}_0) + \eta \mathbf{D}^T \mathbf{D} \mathbf{m}_n] \quad (10)$$

where $\mathbf{G} = \partial \mathbf{g}/\partial \mathbf{m}$ is a matrix of partial derivatives, often referred to as the Fréchet matrix, Fréchet kernel or Jacobian. In practice, it is difficult to compute the derivative of $\mathbf{G}$, so it canbe ignored to

produce the quasi-Newton solution:

$$\delta\mathbf{m}_n = -[\mathbf{G}_n^T \mathbf{C}_d^{-1} \mathbf{G}_n + \varepsilon \mathbf{C}_m^{-1} + \eta \mathbf{D}^T \mathbf{D}]^{-1} [\mathbf{G}_n^T \mathbf{C}_d^{-1} [\mathbf{g}(\mathbf{m}_n) - \mathbf{d}_{\text{obs}}]$$

$$+ \varepsilon \mathbf{C}_m^{-1} (\mathbf{m}_n - \mathbf{m}_0) + \eta \mathbf{D}^T \mathbf{D} \mathbf{m}_n] \tag{11}$$

Both methods require a system of M linear equations to be solved, where M is the number of unknowns.

If instead it is assumed that the relationship $\mathbf{d} = \mathbf{g}(\mathbf{m})$ is linearizable, so that $\delta\mathbf{d} = \mathbf{G}\delta\mathbf{m}$, then the objective function can be written:

$$S(\mathbf{m}) = (\mathbf{G}\delta\mathbf{m} - \delta\mathbf{d})^T \mathbf{C}_d^{-1} (\mathbf{G}\delta\mathbf{m} - \delta\mathbf{d}) + \varepsilon \delta\mathbf{m}^T \mathbf{C}_m^{-1} \delta\mathbf{m}$$

$$+ \eta \delta\mathbf{m}^T \mathbf{D}^T \mathbf{D} \delta\mathbf{m} \tag{12}$$

where the last term on the right hand side of the equation smooths the model perturbations. Setting the derivative of this function to zero yields:

$$\delta\mathbf{m} = [\mathbf{G}^T \mathbf{C}_d^{-1} \mathbf{G} + \varepsilon \mathbf{C}_m^{-1} + \eta \mathbf{D}^T \mathbf{D}]^{-1} \mathbf{G}^T \mathbf{C}_d^{-1} \delta\mathbf{d} \tag{13}$$

When no smoothing is applied and $\mathbf{C}_d$ and $\mathbf{C}_m$ represent known error statistics on the data and prior model respectively, then:

$$\delta\mathbf{m} = [\mathbf{G}^T \mathbf{C}_d^{-1} \mathbf{G} + \mathbf{C}_m^{-1}]^{-1} \mathbf{G}^T \mathbf{C}_d^{-1} \delta\mathbf{d} \tag{14}$$

which is the maximum likelihood solution to the inverse problem or the *stochastic inverse.* Eqs. (11), (13) and (14) are often referred to as damped least square (DLS) solutions, and are the most common class of technique used to solve the inverse problem in seismic tomography (Aki et al., 1977; Thurber, 1983; Eberhart-Phillips, 1986; Farra and Madariaga, 1988; White, 1989; Zhao et al., 1992; Wang and Braile, 1996; Zelt and Barton, 1998; Graeber et al., 2002; Rawlinson et al., 2006b).

All DLS-type schemes ultimately require the solution of a system of M equations, which can be done in various ways. For problems of modest size LU decomposition, Cholesky decomposition or singular value decomposition (SVD) may suffice, but when M is very large, iterative techniques such as the conjugate gradient method (or variants such as LSQR) which take advantage of the sparse nature of linear systems in seismic tomography, may be more effective (Hestenes and Stiefel, 1952; Nolet, 1985; Scales, 1987; VanDecar and Snieder, 1994). Rather than formulate the DLS solution as a set of normal equations, it is also possible to express it directly as a matrix equation:

$$\begin{bmatrix} \mathbf{C}_d^{-1/2} \mathbf{G} \\ \sqrt{\varepsilon} \mathbf{C}_m^{-1/2} \\ \sqrt{\eta} \mathbf{D} \end{bmatrix} \delta\mathbf{m} = \begin{bmatrix} \mathbf{C}_d^{-1/2} \delta\mathbf{d} \\ 0 \\ 0 \end{bmatrix} \tag{15}$$

Application of LSQR or SVD will solve this system in the least-squares sense, which will yield the same solution as Eq. (13).

Rather than minimize $S(\mathbf{m})$ across all M-dimensions of model space, it can be advantageous to restrict the inverse problem to a smaller N-dimensional subspace, where $N < M$. The method of steepest descent is the simplest of these approaches, because it performs a 1D line minimization in the direction specified by the gradient vector $\gamma = \partial S / \partial \mathbf{m}$. Although straightforward to implement and rapid to solve at each iteration, it suffers from poor convergence properties (Rawlinson and Sambridge, 2003b). A more effective approach is conjugate gradients, which was first applied to unconstrained optimization by Fletcher and Reeves (1964). Like steepest descent, each iteration of the technique involves a simple 1D minimization, but in the case of conjugate gradients, the nth iteration locates the minimum in an n-dimensional subspace spanned by the current search direction and all those that precede it. More general subspace techniques, which are not restricted to 1D minimizations at each iteration, have been developed and applied in seismic tomography (Kennett et al., 1988; Sambridge, 1990; Williamson, 1990; Rawlinson et al., 2006b).

4.2.2. Fréchet matrix

All gradient methods require the calculation of the Fréchet matrix $\mathbf{G} = \partial \mathbf{g} / \partial \mathbf{m}$, which describes the rate of change of observables with respect to the model parameters. If the model prediction is made using geometric ray theory, with the required quantity calculated by line integration, then first-order accurate approximations can be readily made. For example, the linearized relationship between traveltime residual and velocity perturbation can be expressed as a simple integral, which can be differentiated to obtain an expression for the rate of change of traveltime with respect to each model parameter. Equivalent expressions can also be obtained for interface and source location parameters (see Rawlinson and Sambridge, 2003b, for more details). Derivatives of other quantities such as attenuation can be obtained in a similar way.

In finite frequency tomography, the situation is more complex, because each observable has a dependence on parameters that lie within a finite volume surrounding the geometric ray. However, as discussed earlier, Dahlen et al. (2000) describe a finite frequency forward solver that only requires a single geometric ray between source and receiver to be computed. The Fréchet kernel is then approximated using a formulation based on paraxial ray theory, which allows information in the neighbourhood of the central ray to be computed. The efficiency of a ray based approach makes it feasible for large tomographic problems (e.g. Montelli et al., 2004), which would otherwise be computationally impractical (using normal mode theory, for example). In addition to computing finite frequency traveltimes, the Born approximation can be used to estimate other seismic observables, including amplitude (Tian et al., 2007a) and time domain waveforms (Panning et al., 2009). This means that it potentially has a role in direct seismic waveform inversion.

The use of paraxial ray theory to approximate sensitivity kernels has had a major impact in the field of seismic tomography. It allows more data to be used to constrain structure, but without the computational overheads of a full waveform inversion. However, more theoretical development is required to address shortcomings in the accuracy of kernels associated with waves that nearly graze a boundary, core diffractions and upper mantle triplications (Nissen-Meyer et al., 2007). These problems arise from the limitations of ray theory in the presence of caustics.

Similar to finite-frequency tomography, one of the main challenges in full waveform tomography is the calculation of the Fréchet kernel. The non-linearity of the inverse problem in full waveform tomography is generally more extreme than in traveltime tomography, so were it not for the prohibitive computational expense, fully non-linear inversion methods would be preferable. One way of circumventing this difficulty is to use a relatively accurate initial model derived from conventional traveltime tomography, and then use a gradient-based technique to refine structural detail by sequentially adding information from low to high frequencies (Sirgue and Pratt, 2004). The complete absence of ray theory in full waveform tomography raises the issue of how to efficiently compute the Fréchet kernel. It turns out that this can be done without explicit calculation of any partial derivative; rather it is derived from a zero-lag correlation of the forward (from the source) and backward (from the receiver) propagated wavefield (Tarantola, 1984; Pratt and Worthington, 1990; Pratt, 1999; Pratt and Shipp, 1999; Sirgue and Pratt, 2004), which amounts to a multiplication of the two wavefields in the frequency domain. The backward propagated or time-reversed wavefield is sometimes referred to as the adjoint wavefield, and consists of the difference between the observed and predicted waveform propagated backward in time from the receiver. The remarkable feature of this approach is that the forward and backward propagated wavefield need only be computed once in order to obtain the sensitivity kernel. In passive source tomography, adjoint methods have only recently been

developed for the computation of regional and global wave sensitivity kernels (e.g. Tromp et al., 2005; Fichtner et al., 2006a,b). Fichtner et al. (2008) provide theoretical background to regional and global full waveform inversion in the frequency domain using adjoint methods to compute sensitivity kernels.

Another approach to computing the full 3D Fréchet kernel was suggested by Nissen-Meyer et al. (2007), who reduce the problem to a series of six independent 2D solutions which may be found using spectral element or other mesh based solvers. This 3D to 2D reduction strategy greatly reduces the computational cost, but is only applicable to spherically symmetric Earth models, and therefore cannot be used in non-linear inversion strategies. Other schemes for computing the full 3D kernel, such as the scattering integral method (Chen et al., 2007), also exist.

Automatic differentiation (AD) is a recently developed and novel approach for computing sensitivity kernels. It generally takes the form of source code to source code translators which generate a program that can be compiled and executed to produce the desired derivatives, hence obviating the need for deriving and hand coding explicit mathematical formulae. The technique essentially exploits the logic of computer codes, in which output (e.g. two point traveltime) can be linked to the input (e.g. velocity grids), thus creating an avenue for directly determining the derivative of one with respect to the other via repeated application of the chain rule. Sambridge et al. (2008) investigate the potential of AD in geophysical inverse problems, including the calculation of traveltime sensitivity to velocity structure. However, it is yet to be used in seismic tomography.
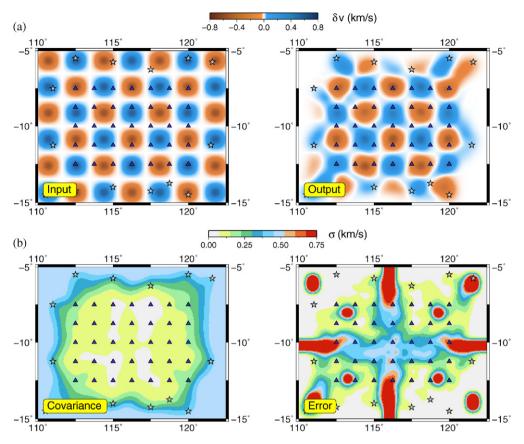
4.3. Fully non-linear inversion

The preceding inversion methods are local in that they only exploit information in limited regions of model space in order to arrive at a solution. The limitation of linear or iterative non-linear techniques is that they are strongly dependent upon accurate initial models, and do not provide robust measures of model uncertainty. The attraction of fully non-linear techniques, which generally rely on exhaustive sampling of model space, is that they can produce an ensemble of data-satisfying models which can be interrogated using statistical techniques to illuminate only those features that are required by the data. The obvious drawback of such an approach is that the computing resources required is many times greater than that needed for a linear or iterative non-linear scheme. For regional or global tomography, which typically constrain a large number of unknowns using massive datasets (e.g. Widiyantoro and van der Hilst, 1997; Bijwaard and Spakman, 2000; Burdick et al., 2008), fully non-linear methods are completely out of the question. However, with the possible exception of full waveform tomography, the inverse problem is not highly non-linear because sub-lithospheric heterogeneity is less severe, and *a priori* information on seismic structure tends to be quite accurate. Therefore the potential advantages of global optimization – even if it was feasible – are less compelling in such circumstances. Crustal or lithospheric scale studies, on the other hand, are more likely to face difficulties with non-linearity and solution non-uniqueness issues due to the presence of strong vertical and lateral heterogeneity that deviates significantly from standard Earth reference models.

Monte Carlo (MC) methods, including simulated annealing and genetic algorithms, are the most common class of fully non-linear search algorithms used in geophysical inverse problems today (Sambridge and Mosegaard, 2001; Mosegaard and Sambridge, 2002). Genetic algorithms use an analog to biological evolution in order to drive the search for new models from an initial pool of randomly generated models, while simulated annealing is based on an analog with physical annealing in thermodynamic systems to guide variations in model parameters. Even with modern cluster comput-

ing, fully non-linear search techniques of this type are limited, at least in the context of seismic tomography, to perhaps a few hundred unknowns at most. A practical alternative that has been used in several studies is to begin with a coarsely parameterized model, apply a non-linear search technique, and then use the solution as a starting model for a local gradient based minimization with a larger number of parameters. The idea behind this hybrid approach is to locate a point in model space sufficiently close to the global minimum solution that locally linearized methods can be successfully used. Applications include Pullammanappallil and Louie (1993) and Boschetti et al. (1996) for the inversion of 2D reflection and refraction traveltimes respectively, and Asad et al. (1999) in the context of 3D local earthquake tomography.

4.4. Analysis of solution robustness

Due to the presence of solution non-uniqueness in all seismic tomography problems, the production of a single data-satisfying model is inadequate, for the central reason that it may contain features that are not required by the data. The lack of a truly robust approach for addressing this issue has spawned the development of a variety of techniques over the years. One of the most common is the synthetic resolution test, in which a heterogeneous input model is used to generate data with the same source–receiver geometry as the observational experiment. The ability of the inversion scheme to recover the input model can then be used as a measure of solution robustness. A commonly implemented example of this approach is the checkerboard resolution test, which uses an input model consisting of an alternating pattern of fast and slow anomalies (Walck, 1988; Glahn and Granet, 1993; Ritsema et al., 1998; Day et al., 2001; Graeber et al., 2002; Rawlinson and Kennett, 2008). Fig. 11a shows the result of a checkerboard resolution test applied to the example in Fig. 1. The recovery of structure within the bounds of the source–receiver array is good, but this is somewhat misleading as it does not capture the strong variability that is actually present (Fig. 11b). The significant non-linearity of the inverse problem is partially responsible for this result, because the path coverage for the checkerboard model differs markedly to that of the actual structure shown in Fig. 1. Synthetic resolution tests are simple to implement and interpret but suffer from several limitations, including that the results can vary according to the input structure used (e.g. Lévêque et al., 1993), as demonstrated in Fig. 11a.

A traditional alternative to synthetic tests comes in the form of posterior covariance and resolution from linear theory (Tarantola, 1987; Menke, 1989), which provides quantitative measures of model uncertainty (Aki et al., 1977; White, 1989; Benz et al., 1992; Wang and Braile, 1996; Graeber and Asch, 1999). In practical tomographic applications, these formal estimates can be difficult to compute and meaningfully interpret, for the following reasons: (1) validity decreases as the non-linearity of the inverse problem increases; (2) inversion of a potentially large $M \times M$ matrix is required; (3) implicit regularization imposed by an assumed model parameterization is not accounted for; (4) *a priori* model covariance is usually poorly known which, coupled with the use of variable damping and smoothing, make the absolute values of posterior uncertainty rather meaningless. The problem of attempting to directly invert large sparse matrices has been overcome in recent times by modifying iterative approaches such as LSQR (Zhang and McMechan, 1995; Yao et al., 1999; Zhang and Thurber, 2007) to approximate the generalized inverse. Fig. 11b shows the posterior covariance matrix associated with the solution model in Fig. 1, together with the actual error. A perfect correlation between the two cannot be expected, because posterior covariance measures how estimates of data uncertainty, together with constraints from prior information, map as uncertainties in the solution model. In this case, the posterior covariance bears some resemblance with the

Fig. 11. Two common techniques for analysing solution robustness applied to the dataset shown in Fig. 1. (a) Synthetic checkerboard test; (b) covariance estimate from linear theory—the plot on the right shows the absolute error (difference between Fig. 1a and d).

actual error, but the smaller wavelength features are not present. The prior model covariance is set to a uniform value of 0.5 km/s, the effect of which is clearly apparent in the posterior covariance (e.g. outside the bounds of the source–receiver array). Within the confines of linearity, one could also analyze solution robustness by identifying the model null space vectors (using SVD for example), and then varying a given solution model only in the null space, such that the data fit remains unchanged. The so-called "nullspace shuttle" proposed by Deal and Nolet (1996) essentially follows this principle, because it allows movement from one model to another without compromising data fit.
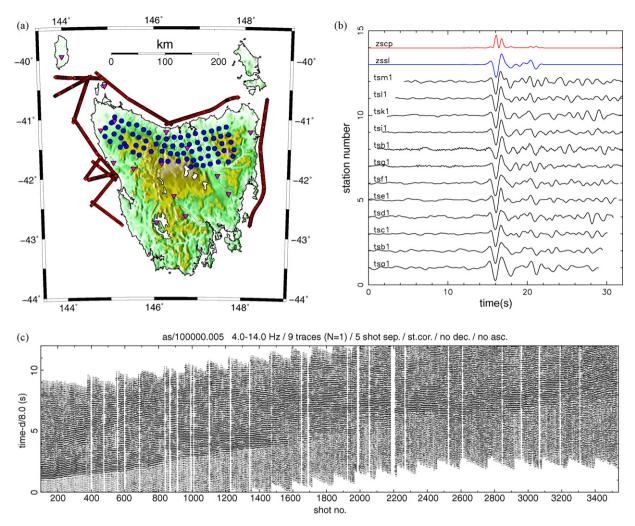
A variety of other methods have been suggested for analyzing solution non-uniqueness; for example, Debayle and Sambridge (2004) estimate the minimum length scale of resolvable structure as a function of location using Voronoi diagrams which are constrained using a quality criterion based on ray density and azimuthal coverage. Statistical methods based on multiple inversion with different components of the dataset such as bootstrapping and jackknifing, have also been used (Lees and Crosson, 1989; Su and Dziewonski, 1997; Zelt, 1999; Gung and Romanowicz, 2004). As pointed out by Nolet et al. (1999), both bootstrapping and jackknifing rely on an overdetermined inverse problem, which is definitely not the case in most tomographic studies e.g. Fig. 1.

The idea of using gradient-based techniques to generate multiple solution models, which can then be assessed for consistent structure, has been around for some time. One possibility is to use a spectrum of initial models, which amounts to starting a local search from multiple points in model space. Vasco et al. (1996) apply cluster analysis to an ensemble of 1075 tomography models generated in this way from cross-hole data. Rawlinson et al. (2008) develop a new technique which exploits information gained from previous solutions to help drive the search for new models. This is done by

adding a feedback or evolution term to the objective function that creates a local maximum at each point in model space occupied by all previous solutions. New models therefore avoid the neighbourhoods of previously generated models, and an ensemble of distinct data-satisfying solutions is produced. Another ensemble approach, discussed earlier in the context of adaptive parameterizations (see end of Section 2.2), is the partition modelling method, first applied to seismic tomography by Bodin and Sambridge (2009). It also produces a set of solution models which can be interrogated for consistent structure, with synthetic tests showing it to be an extremely robust technique.

In many practical tomography problems, ensemble inference techniques are simply not feasible due to computational constraints. When this occurs, no single technique mentioned above can really be considered the "gold standard" for assessing solution robustness, because they all have limitations. Perhaps the best approach in such situations is to apply a variety of methods if available (e.g. covariance/resolution matrices plus synthetic reconstruction) and check their output for consistency, taking into account their strengths and weaknesses. Even if synthetic reconstructions are the only viable option, it is worth doing tests with several different structures to see if resolution is consistent. This would be particularly worthwhile for non-linear problems such as traveltime tomography, where the geometry of ray paths have a structural dependence.

5. Examples

A set of three contrasting tomographic studies is briefly described below in order to put much of the preceding material into a practical context. The first example combines teleseismic and wide-angle traveltimes in a joint inversion for wavespeed and
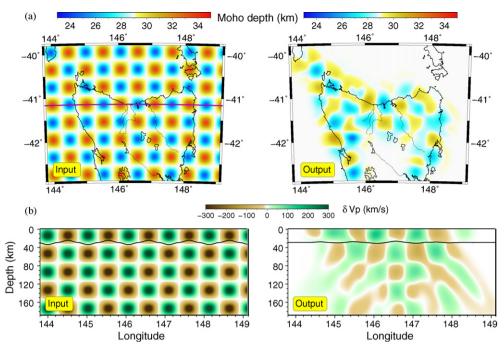
Fig. 12. (a) Location of TIGGER teleseismic array (blue dots), TASGO shot lines (contiguous red diamonds) and TASGO recording stations (magenta triangles). (b) Example of P-wave teleseismic data recorded by a subset of TIGGER array stations, after optimal alignment using adaptive stacking (Rawlinson and Kennett, 2004). Blue trace is the linear stack; red trace is the quadratic stack. (c) Refraction section recorded in NW Tasmania during the TASGO experiment. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

Moho structure beneath Tasmania, southeast Australia. The second example images the anelastic structure beneath the western Pacific Mariana subduction system using both land based and ocean bottom seismometers. The final example uses surface wave tomography to image the 3D shear wavespeed structure of the Australian continent using data from temporary and permanent broadband stations.

5.1. Joint inversion of teleseismic and wide-angle traveltimes

Teleseismic body wave tomography exploits the relative arrival times of global phases recorded by an array of seismometers which overlies the target region. One of the principal assumptions of the method is that lateral heterogeneity outside the local 3D model volume does not significantly contribute to the measured arrival time residuals (difference between observed arrival times and predictions from some reference model). Removing the mean from the residuals on a source-by-source basis eliminates origin time errors, but at the expense of absolute velocity perturbation information. Another problem is that near surface structure is poorly constrained, because rays from distant earthquakes are typically sub-vertical when they arrive at the array. However, near surface structure tends to be strongly heterogeneous, and contributes significantly to the measured arrival time residual. The standard

approach for dealing with this issue is to include station terms as unknowns in the inversion (Frederiksen et al., 1998; Graeber et al., 2002). However, the trade-off between station terms and velocity structure is difficult to resolve, which has led to the more recent trend of incorporating crustal models as prior information (Waldhauser et al., 2002; Lippitsch et al., 2003; Martin and Ritter, 2005; Lei and Zhao, 2007; Rawlinson and Kennett, 2008). A more effective approach is to simultaneously invert all available datasets for a unified model; this is particularly desirable when the datasets contain overlapping constraints.

In this example, both teleseismic and wide-angle traveltime data are incorporated in a joint inversion for the lithospheric structure beneath northern Tasmania (see Rawlinson and Urvoy, 2006, for more details). Fig. 12a shows the geometry of the 72 station TIGGER array which recorded the teleseisms, and the shot lines and land-based stations that recorded the wide-angle data as part of the TASGO experiment. The station spacing of the TIGGER array is approximately 15 km, which means that there is considerable overlap of crossing paths from the two datasets in the lower crust and uppermost mantle. Fig. 12b shows an example of a teleseismic arrival recorded at the TIGGER array, and Fig. 12c shows an example of wide-angle data recorded at a single land-based TASGO station. Refraction and wide-angle reflection phases are picked by hand, and teleseismic arrival time residuals are extracted using

Fig. 13. Checkerboard resolution test for the combined TASGO and TIGGER datasets (a) Moho depth; (b) east-west cross-section showing velocity and Moho depth at 41.2° south (as denoted by magenta line in (a)). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

the adaptive stacking approach of Rawlinson and Kennett (2004), which exploits the coherency of teleseismic waveforms across a dense array.

An iterative non-linear technique is used to invert the travel-time data for both velocity and Moho structure. Seismic structure is represented using regularly spaced velocity and interface grids coupled with cubic B-spline volume elements and surface patches respectively. In this case, the crust and upper mantle are defined as independent layers with the Moho surface representing the interface between the two. The traveltime prediction problem is solved using the Fast Marching Method, a grid based eikonal solver described earlier (see Section 3.3.1). To solve the inverse problem, a subspace inversion method is used (see Section 4.2). The forward and inverse solvers are applied iteratively until both the changes in data fit and model perturbation become insignificant. Fig. 13 shows the result of a synthetic checkerboard test for both velocity and interface structure; the main point here is that the trade-off between velocity and interface geometry appears to be satisfactorily resolved.

A series of slices through the solution model are shown in Fig. 14. Although some smearing is apparent, a number of robust features are clearly present, including a marked thinning of the crust towards the northeast of Tasmania accompanied by elevated velocities. This result is significant, because the traditional view of the Tasmanian lithosphere is that it comprises two separate fragments (East and West Tasmania terranes) that were juxtaposed during the mid-Devonian along the so-called Tamar Fracture System (Williams, 1989). However, combined with evidence from gravity and magnetic data, the results of this study do not support the presence of a crustal scale suture zone. Instead, there appears to be considerably more evidence to support an idea first proposed by Reed (2001) that East and West Tasmania were passively joined as far back as the Ordovician, with subsequent episodes of orogenesis and sediment deposition thickening oceanic East Tasmania and adding it to the pre-existing continental West Tasmania. The thinner crust and elevated wavespeeds observed beneath northeast Tasmania (Fig. 14) are consistent with this model. Other features of the model include elevated wavespeeds beneath the economically

important Mt. Read Volcanics, a Cambrian volcanic belt which hosts sizable deposits of base metals; and some evidence in the upper mantle and crust of fossil subduction (Fig. 14d) associated with the Delamerian-Tyennan subduction system that existed along the proto-Pacific margin of east Gondwana in the Cambrian.

5.2. Attenuation tomography in a subduction zone setting

The depth extent and distribution of temperature anomalies and fluids, including melt and slab-derived volatiles, within a subduction zone mantle wedge are of great importance for understanding dynamics of subduction systems. Geochemical studies at the Mariana subduction system show that melt formation must be influenced by some volatile components that come from the slab (e.g. Pearce et al., 2005), which suggests a need to understand the spatial relationships of melt production and volatile transport regions between different parts of the subduction system. One subsurface imaging tool that can help to identify these spatial variations is seismic attenuation tomography. Experimental studies have shown that temperature (e.g. Jackson et al., 1992), volatiles dissolved in normally anhydrous mantle minerals (Aizawa et al., 2008), and possibly small amounts of melt (Faul et al., 2004) can all have significant effects on seismic attenuation. As such, using seismic attenuation tomography as an imaging tool can provide constraints on thermal anomalies and variations in melt and volatile content.

In this example, local earthquake data is used to obtain t^* attenuation estimates for both P and S waves, which are inverted for P wave attenuation (Q_P^{-1}) and Q_P/Q_S structure (see Pozgay et al., 2009, for more details). Fig. 15a shows the earthquakes and the 20 broadband land and 58 ocean-bottom seismometers of the 2003–2004 Mariana seismic experiment that were used for analysis. The deployment was designed to image the forearc, arc, backarc spreading centre, and far backarc to obtain robust tomographic images of the entire arc–backarc system. For each arrival from a given earthquake, frequency-independent effects are corrected for (e.g. free surface and geometric spreading). Non-negative least squares or NNLS (Lawson and Hanson, 1974) is used to invert the
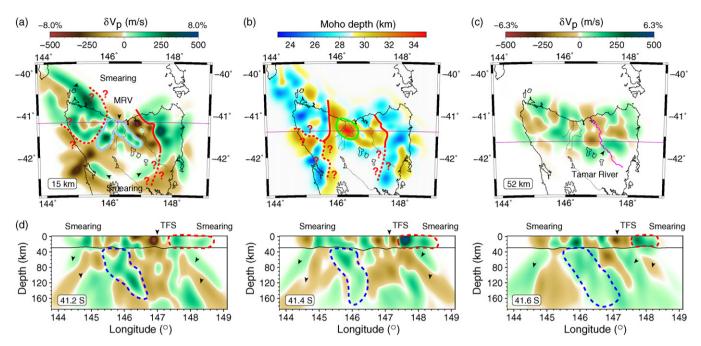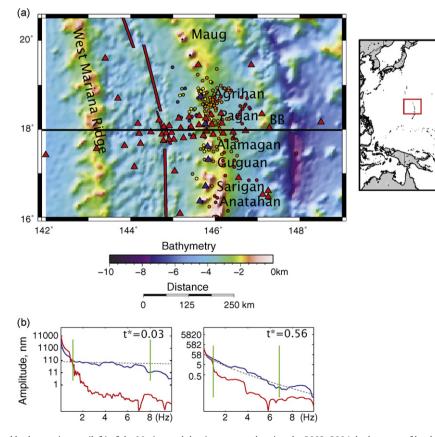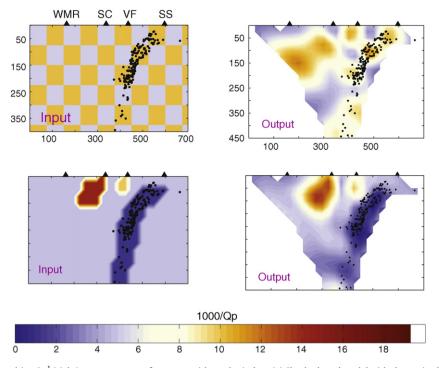
Fig. 14. Results of joint inversion of TASGO and TIGGER traveltime datasets for seismic structure beneath northern Tasmania (from Rawlinson and Urvoy, 2006). (a) Horizontal crustal slice; (b) Moho structure; (c) horizontal mantle slice; (d) three east-west cross-sections with several features highlighted (magenta line in map above each section denotes location). TFS = Tamar Fracture System; MRV = Mt. Read Volcanics. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.) [Figure reproduced from Rawlinson and Urvoy (2006)].

Fig. 15. (a) Location (right) and bathymetric map (left) of the Mariana subduction system showing the 2003–2004 deployment of land stations (blue triangles) and ocean-bottom seismometers (red triangles). The thick black line shows the cross-section used in the tomographic image. Earthquakes used in this study are plotted as small circles and are colour-coded as a function of depth: red <100 km, orange 101–200 km, yellow 201–300 km, green 301–400 km, blue 401–500 km, violet >501 km. The thick red line denotes the backarc spreading center. (b) Example P wave spectra for a forearc OBS (left) and spreading center OBS (right) from an earthquake located at 18.8° 145.7° and 213 km depth on 8 August 2003 at 11:17:48 GMT. Amplitude spectra showing signal spectra (blue) and noise spectra taken from time period immediately prior to the arrival (red). Grey dashed lines show the best-fitting spectral solution. Vertical green bars show spectral limits used in the source parameter and t^* inversion. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.) [Figure reproduced from Pozgay et al. (2009)].

Fig. 16. Input models (left) and resulting Q_P^{-1} (right) output structure from tests with synthetic data. (a) Checkerboard model with alternating blocks of $Q = 200$ and $Q = 100$. (b) Input model resembling realistic asymmetric arc ($Q = 100$) and backarc anomalies ($Q = 70$). See Pozgay et al. (2009) for more models and detailed descriptions. Triangles across the top from left to right correspond to the West Mariana Ridge (WMR), backarc spreading centre (SC), volcanic front (VF), and the serpentinite seamount (SS). Earthquakes delineate the subducting slab. Axes marked in the top row are the same for the bottom row.
[Figure reproduced from Pozgay et al. (2009)].

group of observed displacement amplitude spectra from one event for (1) the path-averaged spectral decay attenuation parameters ($t_{P,S}^*$) for each station (Fig. 15b) and (2) a single corner frequency and seismic moment for the event to correct for the source mechanism effects.

A total of 2900 t^* estimates are inverted for Q_P^{-1} and Q_P/Q_S structure with a model designed with nodes spaced 25 km apart. P and S velocity models determined from the same dataset (Barklage et al., 2006) are used to trace the raypaths using a three-point pseudo-bending method (see Section 3.2). As the tomographic inversion equation is linear for attenuation (see beginning of Section 4), a starting model is not used. A piece-wise joint inversion based on singular value decomposition (see Section 4.2.1) is used in which Q_P^{-1} is first obtained from t_P^* data. Then, as a separate inversion, Q_P/Q_S is computed from t_s^* by incorporating the Q_P^{-1} model parameters into the matrix of partial derivatives. Results of synthetic checkerboard and resolution modelling tests (Fig. 16) (see Section 4.4) show accurate spatial and amplitude retrieval of attenuation parameters down to 450 km depth and beyond the backarc spreading centre.
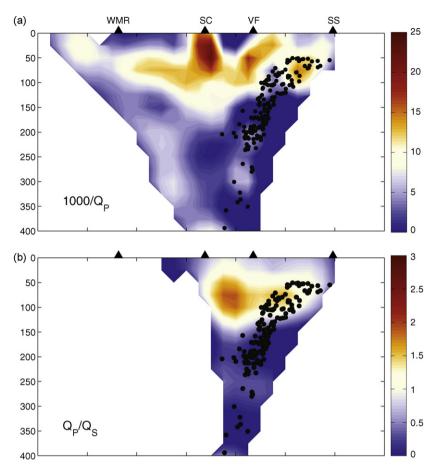
Q_P^{-1} and Q_P/Q_S structures (Fig. 17) show several significant features including low Q_P/Q_S values and a high attenuation region beneath the volcanic arc and a narrow column of very high attenuation directly beneath the backarc spreading centre. The distinct separation at shallow depth of these two high attenuation regions is suggestive of separate melting regimes between the arc and the backarc, an observation also supported by geochemical evidence (Kelley and Plank, submitted for publication); however, the deep connection between the two may be the locus for material transfer at depth, corroborating different geochemical signatures observed between shallow and deep subduction components (Pearce et al., 2005). The narrow column of very high attenuation material beneath the backarc spreading centre is in marked contrast to the only other study to date of attenuation at mantle depths

beneath a spreading centre (Roth et al., 1999), which showed a (relatively) much larger and broader swath of high attenuation beneath the Central Lau Spreading Centre. This striking difference may be indicative of the suggested dominance of a passive upwelling regime at a fast spreading ridge (e.g. Lau) versus active dynamic upwelling at a slow-spreading ridge (e.g. Mariana) (Madge and Sparks, 1997; Parmentier and Phipps Morgan, 1990). As absolute Q^{-1} values (and inferred temperatures) are very high and Q_P/Q_S values are low, the arc and wedge core anomalies are interpreted as regions of high temperature with enhanced attenuation due to hydration and/or melt; the moderate Q^{-1} slab and forearc anomalies as indicative of slab derived fluids and/or large-scale serpentinization; and the columnar-shaped high Q_P^{-1} anomaly directly beneath the backarc spreading center as indicative of a narrow region of dynamic upwelling and melt production beneath the slow-spreading ridge.

5.3. Regional surface wave tomography

The Australian region offers an interesting setting for the comparison of different tomographic inversions. Geologically, the continent is composed of both an old Precambrian shield in the centre and west, and younger Phanerozoic terranes to the east. The eastern margin of the continent has also been affected by Cenozoic volcanism, as recently as 4600 years ago in the Newer Volcanic province in southeast Australia. This range of tectonic settings provides an ideal natural laboratory for studying the variations in seismic velocity associated with the different upper mantle structures.
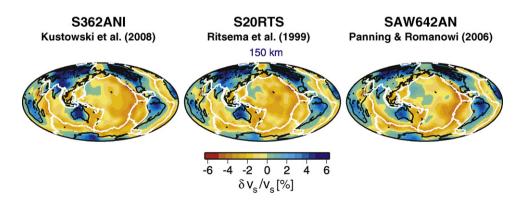
Using combined body and surface wave datasets, recent global models (e.g. Kustowski et al., 2008; Ritsema et al., 1999; Panning and Romanowicz, 2006) clearly show the strong contrast in shear velocity beneath the Australian region. At 150 km depth, the highest velocities are observed beneath central and western Australia,

Fig. 17. P wave (top) and Q_P/Q_S attenuation structures from the SVD inversion projected along the east-west line shown in Fig. 15a. Only nodes that have crossing rays are shown (see Pozgay et al., 2009). Circles are earthquakes used in the study and triangles across the top from left to right correspond to the West Mariana Ridge (WMR), backarc spreading centre (SC), volcanic front (VF), and the serpentinite seamount (SS) (ref. Fig. 15a).
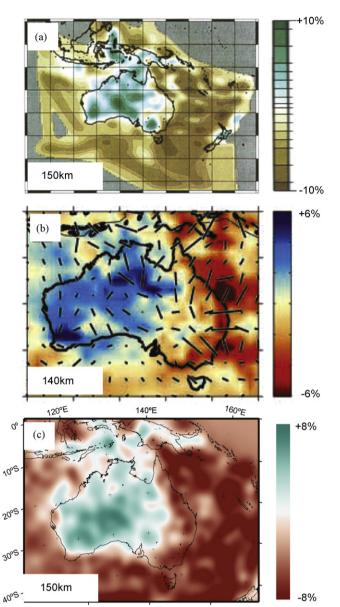[Figure reproduced from Pozgay et al. (2009)].

with lower velocities on the margin of the continent and to the east (see Fig. 18). One of the limitations common to most global models is that the parameterizations used to represent structure are not designed to investigate small-scale velocity variations. Additionally, the small number of seismic stations within Australia that are part of the global seismic networks limits the potential resolution of the models.

Regional surface wave studies are an ideal alternative for investigating the structure of the uppermost mantle. The subduction zones to the north and east generate frequent earthquakes across a broad region, while the mid ocean ridges to the south and east enable data to be incorporated from most azimuthal directions. Although the Australian continent contains only a limited number of permanent seismic stations, the large interior and scarcity of urban areas makes for an ideal location for the deployment of broadband seismometers. The SKIPPY project (van der Hilst et al., 1994) began in 1993, and used a rolling array of broadband instruments to achieve continent-wide coverage with a station spacing of approximately 400 km. Since SKIPPY there have been a number of other temporary networks focused on different parts of Australia.

Fig. 18. Comparison of three global shear velocity models at 150 km depth: S362ANI (Kustowski et al., 2008), S20RTS (Ritsema et al., 1999) and SAW642AN (Panning and Romanowicz, 2006). All three models show the strong contrast between the high velocities observed beneath central and western Australia and the low velocities on the eastern margin of the continent.
[Modified from Kustowski et al. (2008). Copyright 2008 American Geophysical Union. Reproduced by permission of American Geophysical Union].

Fig. 19. Comparison of three shear velocity models from regional surface wave tomography: (a) Debayle and Kennett (2000), (b) Simons et al. (2002), (c) Fishwick and Reading (2008). Each model has been constructed using slightly different datasets, inversion techniques and parameterisations. All models exhibit smaller scale variations in velocity structure compared to the global studies of Fig. 18.

The data from these temporary deployments have been included in a number of surface wave studies (e.g. Zielhuis and van der Hilst, 1996; van der Hilst et al., 1998; Simons et al., 2002; Debayle and Kennett, 2000; Yoshizawa and Kennett, 2004; Fishwick et al., 2005).

Fig. 19 illustrates the shear velocity models at approximately 150 km depth from the studies of Debayle and Kennett (2000), Simons et al. (2002) and Fishwick and Reading (2008). All three inversions use different parameterization schemes: Debayle and Kennett (2000) use the continuous regionalization scheme of Montagner and Nataf (1986), Simons et al. (2002) a discrete regularized inversion with 2° cell sizes, and Fishwick and Reading (2008) spherical B-splines with 2° spacing of knot points (see Section 2 for further discussion on parameterizations). It is, however, difficult to assess the impact of the different parameterization schemes. The path coverage, waveform inversion, and treatment of the third dimension within the tomography will all affect the results (see Simons et al., 2002, for discussion). It does seem likely that the

choice of regularization, or in the case of the continuous regionalization scheme the correlation length, is at least as important as the parameterization.

In contrast to the global studies, the regional surface wave models provide more detailed images of the uppermost mantle. In the recent work of Fishwick et al. (2008) and Fishwick and Reading (2008) over 2600 path-average models have been incorporated into the tomography, with temporary stations in Western Australia giving a significant improvement in path coverage compared to earlier work. Fig. 20 shows three depth slices (75, 150 and 250 km) from the final isotropic model alongside a map showing the main tectonic units. All slices are plotted as velocity perturbations with respect to the global reference model *ak135*. On the eastern margin of the continent, but inland of the continent–ocean transition, low velocities are observed, and a very strong horizontal gradient in velocity continues to depths of around 150 km (Fishwick et al., 2008). There is a strong correlation between the region of low velocity and the location of both high topography and the recent volcanic activity. The transition to the thick lithosphere typical of the Precambrian shield appears to occur as a series of steps, not dissimilar to the structures observed crossing the Trans European Suture Zone (e.g. Shomali and Roberts, 2002; Cotte and Pedersen, 2002; Plomerová et al., 2002). Within the shield region a significant feature of the model is the change from relatively low to relatively high velocities beneath central Australia in the uppermost mantle. This feature is difficult to explain for typical continental geotherms and a constant composition (Fishwick and Reading, 2008). Intriguingly, a similar increase in velocity has recently been suggested to occur beneath a number of cratonic regions (Lebedev et al., 2009; Pedersen et al., 2009). The cause of this anomaly remains uncertain, but seems most likely to represent some compositional variation.

All the surface wave models presented here use great-circle geometric ray paths. Yoshizawa and Kennett (2004) incorporate off-great-circle propagation and a theoretical estimate of the frequency dependent influence zone. In the comparison of models that are constructed with or without these effects, although they observe some changes to the final images, the pattern of velocity anomalies yields analogous features. Most recently, Fichtner et al. (2009) has produced a tomographic model of Australia using full waveform tomography. The similarities between the new work and older models suggests that while the theoretical developments will lead to more complete treatment of seismic data, models carefully constructed using more approximate theory remain useful.

6. Future developments

Over the last three decades or so, seismic tomography has experienced rapid advances on many fronts, including improved techniques for solving the forward and inverse problems; availability of increasingly large volumes of high quality digital data; access to much more powerful computers; and development of new methods for extracting information from data. Given the unpredictable nature of the research horizon – for example, ambient noise tomography emerged with little precursory activity – attempting to forecast the future of seismic tomography is a challenge. Nevertheless, through careful examination of emerging trends, it should be possible to shed some light on where the field might be heading in the next decade or so.

At regional and global scales, one of the main impediments to improving the detail of tomographic models is a lack of good data coverage. The Global Seismographic Network (GSN) comprises over 150 broadband stations and provides free real-time open access to all recorded data. However, the geographic distribution of sites is heavily (and understandably) biased towards land-based stations, with only sparse coverage of ocean basins. This is also true of the combined stations of the International Federation of Digi-
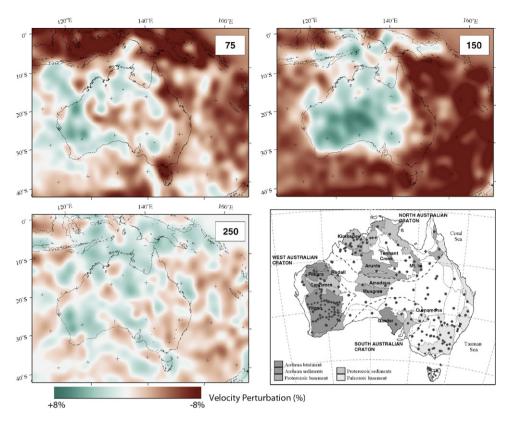
Fig. 20. Images of the isotropic shear velocity models of Fishwick and Reading (2008) at 75 km, 150 km and 250 km depth. The models are plotted as perturbations from the global reference model *ak135* (Kennett et al., 1995) and the same colour scale is used at all depths. A map of the main geological units is also shown for reference.

tal Seismograph Networks (see Fig. 21) or FDSN, which is a global organization of which the GSN is a part. Although new installations are planned for the future, a uniform global coverage of stations is still a long way off. A novel approach to subverting the high cost of *in situ* ocean bottom seismometers was suggested by Simons et al. (2009); they propose the deployment of autonomous floating devices that freely drift in the ocean and record teleseismic events using a hydrophone. These can be cheaply constructed by customizing SOLO floats, which are a popular platform used in oceanography. A group of such instruments, set adrift at judicious locations, could

drastically improve the sampling of the global wavefield in ocean settings.

Temporary array deployments have traditionally played an important role in local and regional scale tomography. The idea of using a rolling array of seismometers to cover a large geographic region has been around for over a decade, with one of the first implementations taking place in Australia with the continent wide SKIPPY experiment (van der Hilst et al., 1994; Zielhuis and van der Hilst, 1996). Today, the largest and most ambitious of these programs is the USArray, which primarily aims to cover continen-
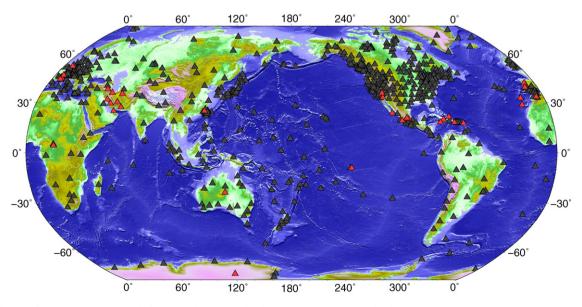

Fig. 21. Distribution of FDSN stations at the end of 2007. Dark grey triangles denote existing stations, and red triangles denote planned stations. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)
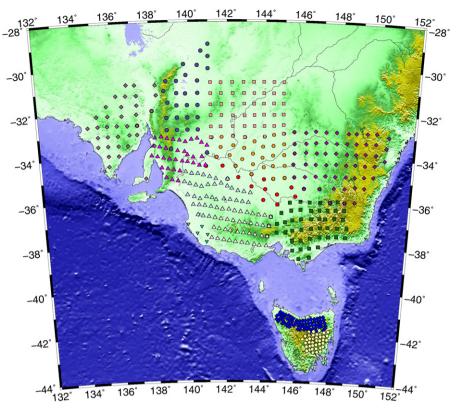
Fig. 22. Cumulative coverage of the WOMBAT rolling seismic array, comprising twelve separate deployments (stations from each are denoted by a unique colour and shape) between 1998 and 2009 in southeast Australia. In each case, deployment periods range between 4 and 14 months.

tal USA with a dense network of portable seismographs over a 15-year period. The transportable array comprises 400 broadband instruments which are installed at approximately 70 km intervals. The bulk of the deployment will be complete by the end of 2013, with coverage of Alaska beginning in 2014. The huge volumes of recorded data are and will continue to be freely available from the IRIS (Incorporated Research Institutions for Seismology) DMC (Data Management Center), and will result in, amongst other things, vastly improved images of the crust and upper mantle beneath the North American continent (e.g. Burdick et al., 2008). In southeast Australia, a rolling seismic array program known as WOMBAT has been in place since 1998, albeit at a much more modest scale. To date, twelve separate array deployments have taken place with a cumulative coverage of over 500 stations at a spacing of between 15 and 50 km (see Fig. 22). The encouraging tomography results obtained so far (Graeber et al., 2002; Rawlinson et al., 2006a,b; Rawlinson and Urvoy, 2006; Clifford et al., 2008; Rawlinson and Kennett, 2008) mean that this program is set to continue, with the ultimate goal being the high density coverage of eastern Australia.

Although the main focus areas for future development in seismic tomography are finite frequency, full waveform and ambient noise tomography, advances in other areas will also continue. For example, one potentially productive area of development is in the joint inversion of overlapping or complementary datasets. This might take the form of active and passive datasets from a similar geographic region; early work in this area was done by Thurber (1983) and Ankeny et al. (1986), who included refraction traveltimes from explosive sources in a LET study. Other combinations include local earthquake, explosive source/airgun traveltimes (Parsons and Zoback, 1997; Wagner et al., 2007); local earthquake and teleseismic traveltimes (Sato et al., 1996); coincident reflection and wide-angle (Wang and Braile, 1996; McCaughey and Singh, 1997); and teleseismic and wide-angle (Rawlinson and Urvoy, 2006). In fact any combination is possible, and will yield better results than

separate inversions of the individual datasets provided that they constrain at least a common subset of parameters. Another case in point is joint inversion of surface waves and teleseismic receiver functions (Özalaybey et al., 1997; Du and Foulger, 1999; Julia et al., 2003; Tkalčić et al., 2006); long period surface waves poorly constrain the crust, while receiver functions provide detailed information on crust and uppermost mantle structure, but do not resolve absolute shear wavespeed very accurately. Conceivably, one could invert any number of overlapping datasets, including body wave traveltimes for any source–receiver geometry, surface waveforms and ambient noise cross-correlations.

The idea of jointly inverting multiple datasets need not be limited to seismic data alone; indeed, it is possible to integrate other classes of geophysical datasets including gravity and magnetic anomalies. For example, simultaneous inversion of traveltimes and gravity data has been carried out by a number of authors (e.g. Lees and VanDecar, 1991; Roy et al., 2005), although some empirical relationship between velocity and density is usually invoked to allow common parameters to be constrained by both datasets. Combining multiple datasets of different types is more widespread in exploration seismology than other areas of seismology, with various techniques proposed for the joint inversion of seismic, gravity and electromagnetic data in existence (e.g. Colombo and De Stefano, 2007). The main challenge is to identify reasonable cross-parameter constraints, which can take many forms, including empirical, physical and statistical. Another approach to joint imaging that is worth mentioning are recent attempts to directly invert seismic and other data for mantle composition and thermal state (Cammarano et al., 2005; Khan et al., 2008). For seismic data, this can be done by using results from mineral physics experiments that analyze the seismic properties of different rock types. Although only applied to 1D models so far, the technique has the major attraction of bypassing the difficult problem of attempting (usually qualitatively) to associate seismic properties (such as com-

pressional wavespeed) with physical and chemical properties of the Earth (e.g. temperature and composition). That said, it is also possible to use a two stage procedure to go from seismic observables to seismic properties and then to physical and chemical properties. Shito et al. (2006) develop a scheme for inverting velocity and attenuation tomographic models for the 3D distribution of temperature, water content and other parameters (e.g. major element geochemistry, partial melting), again invoking results from mineral physics observations. They apply the technique to data from the Izu-Bonin subduction zone and show that lateral variations in seismic properties can be largely explained by variations in temperature and water content.

Geometric ray tracing, for many decades the stalwart of the data prediction step in seismic tomography, may be at the cross-roads in terms of new developments. Over the last few decades, exhaustive research has been carried out in an attempt to find efficient methods to solve the two point (source–receiver path) problem. In recent times, grid-based methods including eikonal solvers and SPR have started to supercede ray tracing in some areas. Now, with finite frequency and full waveform tomography emerging on the scene, new advances in ray tracing appear to be on the wane. This is not due to redundancy, because two point rays are usually still required in both fields; in the case of finite frequency tomography, the central ray is needed for the paraxial approximation, and in full waveform tomography, an initial model generated using traditional traveltime tomography (for example) is often used. One explanation may be that with the power of modern computers, many of the techniques that have been developed in the past are now efficient enough to tackle large problems. There is one area that is still experiencing rapid development—that of multi-arrival techniques, as discussed in Section 3.4. Ray based wavefront construction techniques have just started to be used in practical seismic imaging problems, and grid-based techniques are beginning to become viable alternatives. The question of how to best exploit them in practical seismic tomography applications still remains, however.

Ambient noise tomography, which has only been around for half a decade or so (Shapiro et al., 2005), has dramatically changed the seismic imaging landscape. The ability to largely control data coverage through the geographical distribution of a seismic array, and the fact that ambient noise information is independent of and often complimentary to information obtained from deterministic sources, are two of its major attractions. Rapid adoption of this technique throughout the seismic imaging community is set to continue, and new developments will improve its usefulness. For example, most studies produce maps of Rayleigh wave group wavespeed at different periods (e.g. Shapiro et al., 2005; Saygin and Kennett, 2009), but 3D shear wavespeed is more conducive to interpretation, and can be produced using methods akin to those encountered in traditional surface wave tomography. Finite frequency tomography has now become an established technique, and its application is likely to escalate over the next decade, as more and more high quality data become available. Full waveform tomography, on the other hand is still in its infancy, but there is little doubt that it holds the potential to vastly improve on the present generation of seismic tomography images. Although current computing power is barely adequate to solve realistic problems, expected increases based on recent history suggest that over the next decade or so, full waveform tomography will begin to emerge as a powerful tool for imaging the earth at local, regional and global scales.

Acknowledgements

References

Abers, G.A., Roecker, S.W., 1991. Deep structure of an arc-continent collision; earthquake relocation and inversion for upper mantle P and S wave velocities beneath Papua New Guinea. J. Geophys. Res. 96, 6379–6401.

Abt, D.L., Fischer, K.M., 2008. Resolving three-dimensional anisotropic structure with shear wave splitting tomography. Geophys. J. Int. 173, 859–886.

Achauer, U., 1994. New ideas on the Kenya rift based on the inversion of the combined dataset of the 1985 and 1989/90 seismic tomography experiments. Tectonophysics 236, 305–329.

Afnimar, Koketsu, K., 2000. Finite difference traveltime calculation for head waves travelling along an irregular interface. Geophys. J. Int. 143, 729–734.

Aizawa, Y., Barnhoorn, A., Faul, U.H., Fitz Gerald, J.D., Jackson, I., Kovcs, I., 2008. Seismic properties of Anita Bay dunite: an exploratory study of the influence of water. J. Petrol. 49, 841–855.

Aki, K., Christoffersson, A., Husebye, E.S., 1977. Determination of the three-dimensional seismic structure of the lithosphere. J. Geophys. Res. 82, 277–296.

Aki, K., Lee, W.H.K., 1976. Determination of the three-dimensional velocity anomalies under a seismic array using first P arrival times from local earthquakes 1. A homogeneous intial model. J. Geophys. Res. 81, 4381–4399.

Amirbekyan, A., Michel, V., Simons, F.J., 2008. Parametrizing surface wave tomoraphic models with harminic spherical splines. Geophys. J. Int. 174, 617–628.

Ankeny, L.A., Braile, L.W., Olsen, K.H., 1986. Upper crustal structure beneath the Jemez Mountains volcanic field, New Mexico, determined by three-dimensional simultaneous inversion of seismic refraction and earthquake data. J. Geophys. Res. 91, 6188–6198.

Asad, A.M., Pullammanappallil, S.K., Anooshehpoor, A., Louie, J.N., 1999. Inversion of travel-time data for earthquake locations and three-dimensional velocity structure in the Eureka Valley area, eastern California. Bull. Seismol. Soc. Am. 89, 796–810.

Backus, G., Gilbert, F., 1969. Constructing P-velocity models to fit restricted sets of travel-time data. Bull. Seismol. Soc. Am. 59, 1407–1414.

Backus, G.E., Gilbert, J.F., 1968. The resolving power of gross earth data. Geophys. J. Royal Astr. Soc. 16, 169–205.

Bai, C.-Y., 2005. 3-D multi-step travel time tomography: imaging the local, deep velocity structure of Rabaul volcano, Papua New Guinea. Phys. Earth Planet. Inter. 151, 259–275.

Barklage, M.E., Conder, J.A., Wiens, D.A., Shore, P.J., Shiobara, H., Sugioka, H., Zhang, H., 2006. 3-D seismic tomography of the Mariana mantle wedge from the 2003–2004 passive component of the Mariana subduction factory imaging experiment. EOS Trans. AGU, Fall Meet. Suppl. 87 (52), T23C–0506.

Benamou, J.D., 1996. Big ray tracing: multivalued travel time field computation using viscosity solutions of the Eikonal equation. J. Comp. Phys. 128, 463–474.

Benamou, J.D., 1999. Direct computation of multivalued phase space solutions for Hamilton–Jacobi equations. Comm. Pure Appl. Math. 52, 1443–1475.

Benson, G.D., Ritzwoller, M.H., Shapiro, N.M., 2008. Broad-band ambient noise surface wave tomography across the United States. J. Geophys. Res. 113, doi:10.1029/2007JB005248.

Benson, G.D., Ritzwoller, M.H., Yang, Y., 2009. A 3D shear velocity model of the crust and uppermost mantle beneath the United States from ambient seismic noise. Geophys. J. Int. 177, 1177–1196.

Benz, H.M., Smith, R.B., 1984. Simultaneous inversion for lateral velocity variations and hypocenters in the Yellowstone region using earthquake and refraction data. J. Geophys. Res. 89, 1208–1220.

Benz, H.M., Zandt, G., Oppenheimer, D.H., 1992. Lithospheric structure of northern California from teleseismic images of the upper mantle. J. Geophys. Res. 97, 4791–4807.

Bhattacharyya, J., Masters, G., Shearer, P., 1996. Global lateral variations of shear wave attenutation in the upper mantle. J. Geophys. Res. 101, 22273–22289.

Bijwaard, H., Spakman, W., 2000. Non-linear global P-wave tomography by iterated linearized inversion. Geophys. J. Int. 141, 71–82.

Bijwaard, H., Spakman, W., Engdahl, E.R., 1998. Closing the gap between regional and global travel time tomography. J. Geophys. Res. 103, 30,055–30,078.

Billien, M., Lèvêque, J.J., 2000. Global maps of Rayleigh wave attenuation for periods between 40 and 150 seconds. Geophys. Res. Lett. 27, 3619–3622.

Bishop, T.P., Bube, K.P., Cutler, R.T., Langan, R.T., Love, P.L., Resnick, J.R., Shuey, R.T., Spindler, D.A., Wyld, H.W., 1985. Tomographic determination of velocity and depth in laterally varying media. Geophysics 50, 903–923.

Bleibinhaus, F., Gebrande, H., 2006. Crustal structure of the Eastern Alps along the TRANSALP profile from wide-angle seismic tomography. Tectonophysics 414, 51–69.

Blundell, C.A., 1993. Resolution analysis of seismic P-wave velocity estimates using reflection tomographic inversion. Ph.D. Thesis, Monash University.

Bodin, T., Sambridge, M., 2009. A self-parametrizing partition model approach to tomographic inverse problems. Inverse Probl. 25, doi:10.1088/0266-5611/25/5/055009.

Böhm, G., Galuppo, P., Vesnaver, A., 2000. 3D adaptive tomography using Delauney triangles and Voronoi polygons. Geophys. Prospect. 48, 723–744.

Bois, P., La Porte, M., Lavergne, M., Thomas, G., 1971. Essai de détermination automatique des vitesses sismiques par mesures entre puits. Geophys. Prospect. 19, 42–83.

Bois, P., La Porte, M., Lavergne, M., Thomas, G., 1972. Well-to-well seismic measurements. Geophysics 37, 471–480.

Boschetti, F., Dentith, M.K., List, R.D., 1996. Inversion of seismic refraction data using genetic algorithms. Geophysics 61, 1715–1727.

Boschi, L., Dziewonski, A.M., 1999. 'High' and 'low' resolution images of the Earth's mantle—implications of different approaches to tomographic modeling. J. Geophys. Res. 104, 25567–25594.

Boschi, L., Dziewonski, A.M., 2000. Whole earth tomography from delay times of P, PcP, and PKP phases: lateral heterogeneities in the outer core or radial anisotropy in the mantle? J. Geophys. Res. 105, 13675–13696.

Bregman, N.D., Bailey, R.C., Chapmans, C.H., 1989. Crosshole seismic tomography. Geophysics 54, 200–215.

Brenders, A.J., Pratt, R.G., 2007. Efficient waveform tomography for litho-spheric imaging: implications for realistic, 2-D acquisition geometries and low frequency data. Geophys. J. Int. 168, 152–170.

Bruneton, M., Pederson, H.A., Farra, V., et al., 2004. Complex lithospheric structure under the central Baltic Shield from surface wave tomography. J. Geophys. Res. 109, B10303, doi:10.1029/2003JB002947.

Bulant, P., 1999. Two-point ray-tracing and controlled initial-value ray-tracing in 3-D heterogeneous block structures. J. Seismic Explor. 8, 57–75.

Burdick, S., Li, C., Martynov, V., Cox, T., Eakins, J., Astiz, L., Vernon, F.L., Pavlis, G.L., Van der Hilst, R.D., 2008. Upper mantle heterogeneity beneath North America from travel time tomography with global and USArray transportable array data. Seismol. Res. Lett. 79, 384–392.

Buske, S., Kästner, U., 2004. Efficient and accurate computation of seismic traveltime and amplitudes. Geophys. Prospect. 52, 313–322.

Calvet, M., Chevrot, S., 2005. Travel time sensitivity kernels for PKP phases in the mantle. Phys. Earth Planet. Inter. 153, 21–31.

Cammarano, F., Goes, S., Deuss, A., Giardini, D., 2005. Is a pyrolitic adia-batic mantle compatible with seismic data? Earth Planet. Sci. Lett. 232, 227–243.

Campillo, M., Paul, A., 2003. Long-range correlations in the diffuse seismic coda. Science 299, 547–549.

Cao, S., Greenhalgh, S., 1994. Finite-difference solution of the eikonal equation using an efficient, first-arrival, wavefront tracking scheme. Geophysics 59, 632–643.

Carcione, J.M., 1994. The wave equation in generalized coordinates. Geophysics 59, 1911–1919.

Cassell, B.R., 1982. A method for calculating synthetic seismograms in laterally varying media. Geophys. J. Royal Astr. Soc. 69, 339–354.

Červený, V., 1987. Ray tracing algorithms in three-dimensional laterally varying layered structures. In: Nolet, G. (Ed.), Seismic Tomography: With Applications in Global Seismology and Exploration Geophysics. D. Reidel, Dordrecht, pp. 99–133.

Červený, V., 2001. Seismic Ray Theory. Cambridge University Press, Cambridge.

Červený, V., Firbas, P., 1984. Numerical modelling and inversion of travel times of seismic body waves in inhomogeneous anisotropic media. Geophys. J. Royal Astr. Soc. 76, 41–51.

Červený, V., Klimes, L., Pšenčik, I., 2007. Seismic ray method: recent developments. Adv. Geophys. 48, 1–126.

Červený, V., Pšenčik, I., 1983. Gaussian beams and paraxial ray approximation in three-dimensional elastic inhomogeneous media. Geophys. J. 53, 1–15.

Chapman, C., 2004. Fundamentals of Seismic Wave Propagation. Cambridge University Press, Cambridge.

Chapman, C.H., Drummond, R., 1982. Body-wave seismograms in inhomogeneous media using Maslov asymptotic theory. Bull. Seismol. Soc. Am. 72, S277–S317.

Chen, P., Jordan, T.H., Zhao, L., 2007. Full three-dimensional tomography: a comparison between the scattering-integral and adjoint-wavefield methods. Geophys. J. Int. 170, 175–181.

Cheng, L.-T., 2007. Efficient level set methods for constructing wavefronts in three spatial dimensions. J. Comp. Phys. 226, 2250–2270.

Cheng, N., House, L., 1996. Minimum traveltime calculations in 3-D graph theory. Geophysics 61, 1895–1898.

Chevrot, S., 2006. Finite-frequency vectorial tomography: a new method for high-resolution imaging of upper mantle anisotropy. Geophys. J. Int. 165, 641–657.

Chevrot, S., Zhao, L., 2007. Multiscale finite-frequency Rayleigh wave tomography of the Kaapvaal craton. Geophys. J. Int. 169, 201–215.

Chiao, L.-Y., Kuo, B.-Y., 2001. Multiscale seismic tomography. Geophys. J. Int. 145, 517–527.

Chiu, S.K.L., Kanasewich, E.R., Phadke, S., 1986. Three-dimensional determination of structure and velocity by seismic tomography. Geophysics 51, 1559–1571.

Chou, C.W., Booker, J.R., 1979. A Backus–Gilbert approach to inversion of travel time data for three-dimensional velocity structure. Geophys. J. Royal Astr. Soc. 59, 325–344.

Clifford, P., Greenhalgh, S., Houseman, G., Graeber, F., 2008. 3-D seismic tomography of the Adelaide fold belt. Geophys. J. Int. 172, 167–186.

Clowes, R.M., Zelt, C.A., Amor, J.R., Ellis, R.M., 1995. Lithospheric structure in the southern Canadian Cordillera from a network of seismic refraction lines. Can. J. Earth Sci. 32, 1485–1513.

Colombo, D., De Stefano, M., 2007. Geophysical modeling via simultaneous joint inversion of seismic, gravity, and electromagnetic data: application to prestack depth imaging. Leading Edge 28, 326–331.

Constable, S.C., Parker, R.L., Constable, C.G., 1987. Occam's inversion: a practical algorithm for generating smooth models from electromagnetic sounding data. Geophysics 52, 289–300.

Cotte, N., Pedersen, H.A., TOR Working Group, 2002. Sharp contrast in lithospheric structure across the Sorgenfrei–Tornquist zone as inferred by Rayleigh wave analysis of TOR1 project data. Tectonophysics 360, 75–88.

Curtis, A., Snieder, R., 1997. Reconditioning inverse problems using the genetic algorithm and revised parameterization. Geophysics 62, 1524–1532.

Dahlen, F.A., Hang, S.H., Nolet, G., 2000. Fréchet kernels for finite frequency travel times. I. Theory. Geophys. J. Int. 141, 157–174.

Dahlen, F.A., Nolet, G., 2005. Comment on 'On sensitivity kernels for 'wave equation' transmission tomography' by de Hoop and van der Hilst. Geophys. J. Int. 163, 949–951.

Dalton, C.A., Ekström, G., 2006. Global models of surface-wave attenuation. J. Geophys. Res. 111, doi:10.1029/2005JB003997.

Dalton, C.A., Ekström, G., Dziewonski, A.M., 2008. The global attenuation structure of the upper mantle. J. Geophys. Res. 113, doi:10.1029/2007JB005429.

Danesi, S., Morelli, A., 2000. Group velocity of Rayleigh waves in the Antarctic region. Phys. Earth Planet. Inter. 122, 55–66.

Darbyshire, F.A., Bjarnason, I.J., White, R.S., Florenz, O.G., 1998. Crustal structure above the Iceland mantle plume, imaged by the ICEMELT refraction profile. Geophys. J. Int. 135, 1131–1149.

Darbyshire, F.A., Lebedev, S., 2009. Rayleigh wave phase-velocity heterogeneity and multilayered azimuthal anisotropy of the Superior Craton, Ontario. Geophys. J. Int. 176, 215–234.

Day, A.J., Peirce, C., Sinha, M.C., 2001. Three-dimensional crustal structure and magma chamber geometry at the intermediate-spreading, back-arc Valu Fa Ridge, Lau Basin—results of a wide-angle seismic tomographic inversion. Geophys. J. Int. 146, 31–52.

de Hoop, M.V., van der Hilst, R.D., 2005a. On sensitivity kernels for "wave equation" tomography. Geophys. J. Int. 160, 621–633.

de Hoop, M.V., van der Hilst, R.D., 2005b. Reply to comment by F.A. Dahlen and G. Nolet on: "On sensitivity kernels for wave-equation transmission tomography". Geophys. J. Int. 163, 952–955.

de Hoop, M.V., van der Hilst, R.D., Shen, P., 2006. Wave-equation reflection tomography: annihilators and sensitivity kernels. Geophys. J. Int. 167, 1332–1352.

Deal, M.M., Nolet, G., 1996. Nullspace shuttles. Geophys. J. Int. 124, 372–380.

Debayle, E., 1999. SV-wave azimuthal anisotropy in the Australian upper mantle: preliminary results from automated Rayleigh waveform inversion. Geophys. J. Int. 137, 747–754.

Debayle, E., Kennett, B., Priestley, K., 2005. Global azimuthal seismic anisotropy and the unique plate-motion deformation of Australia. Nature 433, 509–512.

Debayle, E., Kennett, B.L.N., 2000. The Australian continental upper mantle: structure and deformation inferred from surface waves. J. Geophys. Res. 105, 25423–25450.

Debayle, E., Kennett, B.L.N., 2003. Surface wave studies of the Australian region. In: Hillis, R.R., Müiller, R.D. (Eds.), The Evolution and Dynamics of the Australian Plate. Special Publication. Geological Society of Australia and America, pp. 25–40.

Debayle, E., Sambridge, M., 2004. Inversion of massive surface wave data sets: model construction and resolution assessment. J. Geophys. Res. 109, doi:10.1029/2003JB002652.

Du, Z.J., Foulger, G.R., 1999. The crustal structure beneath the northwest fjords, Iceland, from receiver functions and surface waves. Geophys. J. Int. 139, 419–432.

Dueker, K., Humphreys, E., Biasi, G., 1993. Teleseismic imaging of the western United States upper mantle structure using the simultaneous iterative reconstruction technique. In: Iyer, H.M., Hirahara, K. (Eds.), Seismic Tomography: Theory and Practice. Chapman & Hall, London, pp. 265–298.

Dziewonski, A.M., Hager, B.H., O'Connell, R.J., 1977. Large-scale heterogeneities in the lower mantle. J. Geophys. Res. 82, 239–255.

Dziewonski, A.M., Woodhouse, J.H., 1987. Global images of the earth's interior. Science 236, 37–48.

Eberhart-Phillips, D., 1986. Three-dimensional velocity structure in northern California coast ranges from inversion of local earthquake arrival times. Bull. Seismol. Soc. Am. 76, 1025–1052.

Eberhart-Phillips, D., 1990. Three-dimensional P and S velocity structure in the Coalinga Region, California. J. Geophys. Res. 95, 15,343–15,363.

Eberhart-Phillips, D., Henderson, C.M., 2004. Inclusing anisotropy in 3-D velocity inversion and application to Marlborough, New Zealand. Geophys. J. Int. 156, 237–254.

Eberhart-Phillips, D., Michael, A.J., 1993. Three-dimensional velocity structure, seismicity, and fault structure in the Parkfield Region, central California. J. Geophys. Res. 98, 15737–15758.

Eberhart-Phillips, D., Reyners, M., 1997. Continental subduction and three-dimensional crustal structure: the northern South Island, New Zealand. J. Geophys. Res. 102, 11848–11861.

Ekström, G., Tromp, J., Larson, E.W.F., 1997. Measurements and global models of surface wave propagation. J. Geophys. Res. 102, 8137–8157.

Engquist, B., Runborg, O., Tornberg, A.-K., 2002. High-frequency wave propagation by the segment projection method. J. Comp. Phys. 178, 373–390.

Faccioli, E., Maggio, F., Paolucci, R., Quarteroni, A., 1997. 2D and 3D elastic wave propagation by a pseudo-spectral domain decomposition method. J. Seismol. 1, 237–251.

Farra, V., Madariaga, R., 1987. Seismic waveform modelling in heterogeneous media by ray perturbation theory. J. Geophys. Res. 92, 2697–2712.

Farra, V., Madariaga, R., 1988. Non-linear reflection tomography. Geophys. J. 95, 135–147.

Faul, U.H., Fitz Gerald, J.D., Jackson, I., 2004. Shear wave attenuation and dispersion in melt-bearing olivine polycrystals. 2. Microstructural interpretation and seismological implications. J. Geophys. Res. 109, doi:10.1029/2003JB002407.

Favier, N., Chevrot, S., 2003. Sensitivity kernels for shear wave splitting in transverse isotropic media. Geophys. J. Int. 153, 213–228.

Fichtner, A., Bunge, H.P., Igel, H., 2006a. The adjoint method in seismology. I. Theory. Phys. Earth Planet. Inter. 157, 86–104.

Fichtner, A., Bunge, H.P., Igel, H., 2006b. The adjoint method in seismology. II. Applications: traveltimes and sensitivity functionals. Phys. Earth Planet. Inter. 157, 105–123.

Fichtner, A., Kennett, B.L.N., Igel, H., Bunge, H.P., 2008. Theoretical background for continental and global scale full-waveform inversion in the time-frequency domain. Geophys. J. Int. 175, 665–685.

Fichtner, A., Kennett, B. L. N., Igel, H. and Bunge, H.-P., 2009. Full seismic waveform tomography for upper-mantle structure in the Australasian region using adjoint methods. Geophysical Journal International, doi:10.1111/j.1365-246X.2009.04368.x, in press.

Fischer, R., Lees, J.M., 1993. Shortest path ray tracing with sparse graphs. Geophysics 58, 987–996.

Fishwick, S., Heintz, M., Kennett, B.L.N., Reading, A.M., Yoshizawa, K., 2008. Steps in lithospheric thickness within eastern Australia, evidence from surface wave tomography. Tectonics 27, doi:10.1029/2007TC002116.

Fishwick, S., Kennett, B.L.N., Reading, A.M., 2005. Contrasts in lithospheric structure within the Australian craton—insights from surface wave tomography. Earth Planet. Sci. Lett. 231, 163–176.

Fishwick, S., Reading, A.M., 2008. Anomalous lithosphere beneath the Proterozoic of western and central Australia: a record of continental collision and intraplate deformation? Precam. Res. 166, 111–121.

Fletcher, R., Reeves, C.M., 1964. Function minimization by conjugate gradients. Comput. J. 7, 149–154.

Fomel, S., Sethian, J.A., 2002. Fast-phase space computation of multiple arrivals. Proc. Natl. Acad. Sci. 99, 7329–7334.

Forsyth, D.W., Li, A., 2005. Array analysis of two-dimensional variations in surface wave phase velocity and azimuthal anisotropy in the presence of multipathing interference. In: Levander, A., Nolet, G. (Eds.), Seismic Earth: Array Analysis of Broadband Seismograms. AGU Geophysical Monograph Series, Washington, DC, pp. 81–97.

Frankel, A., Vidale, J., 1992. A three dimensional simulation of seismic waves in the Santa Clara Valley, California, from a Loma Prieta aftershock. Bull. Seismol. Soc. Am. 82, 2045–2074.

Frederiksen, A.W., Bostock, M.G., VanDecar, J.C., Cassidy, J.F., 1998. Seismic structure of the upper mantle beneath the northern Canadian Cordillera from teleseismic travel-time inversion. Tectonophysics 294, 43–55.

Friederich, W., 1999. Propagation of seismic shear and surface waves in a laterally heterogeneous mantle by multiple forward scattering. Geophys. J. Int. 136, 180–204.

Friederich, W., 2003. The S-velocity structure of the East Asian mantle from inversion of shear and surface waveforms. Geophys. J. Int. 153, 88–102.

Friederich, W., Wielandt, E., 1995. Interpretation of seismic surface waves in regional networks: joint estimation of wavefield geometry and local phase velocity. Method and numerical test. Geophys. J. Int. 120, 731–744.

Fukao, Y., Obayashi, M., Inoue, H., Nebai, M., 1992. Subducting slabs stagnant in the mantle. J. Geophys. Res. 97, 4809–4822.

Furumura, T., Kennett, B.L.N., Furumura, M., 1998. Seismic wavefield calculation for laterally heterogeneous whole earth models using the pseu-dospectral method. Geophys. J. Int. 135, 845–860.

Glahn, A., Granet, M., 1993. Southern Rhine Graben: small-wavelength to-mographic study and implications for the dynamic evolution of the graben. Geophys. J. Int. 113, 399–418.

Graeber, F.M., Asch, G., 1999. Three-dimensional models of P wave velocity and P-to-S velocity ratio in the southern central Andes by simultaneous inversion of local earthquake data. J. Geophys. Res. 104, 20237–20256.

Graeber, F.M., Houseman, G.A., Greenhalgh, S.A., 2002. Regional teleseis-mic tomography of the western Lachlan Orogen and the Newer Volcanic Province, southeast Australia. Geophys. J. Int. 149, 249–266.

Grand, S.P., van der Hilst, R.D., Widiyantoro, S., 1997. Global seismic tomography: a snapshot of convection in the Earth. GSA Today 7, 1–7.

Granet, M., Trampert, J., 1989. Large-scale P-velocity structures in the Euro-Mediterranean area. Geophys. J. Int. 99, 583–594.

Graves, R.W., 1996. Simulating seismic wave propagation in 3D elastic media using staggered-grid finite differences. Bull. Seismol. Soc. Am. 86, 1091–1106.

Guiziou, J.L., Mallet, J.L., Madariaga, R., 1996. 3-D seismic reflection tomography on top of the GOCAD depth modeler. Geophysics 61, 1499–1510.

Gung, Y., Romanowicz, B., 2004. Q tomography of the upper mantle using three-component long-period waveforms. Geophys. J. Int. 157, 813–830.

Hammer, P.T.C., Dorman, L.M., Hildebrand, J.A., Cornuelle, B.D., 1994. Jasper Seamount structure: seafloor seismic refraction tomography. J. Geophys. Res. 99, 6731–6752.

Hauser, J., Sambridge, M., Rawlinson, N., 2008. Multiarrival wave-front tracking and its applications. Geochem. Geophys. Geosyst. 9, doi:10.1029/2008GC002069.

Heintz, M., Debayle, E., Vauchez, A., 2005. Upper mantle structure of the South American continent and neighboring oceans from surface wave tomography. Tectonophysics 406, 115–139.

Hestenes, M., Stiefel, E., 1952. Methods of conjugate gradients for solving linear systems. Nat. Bur. Stand. J. Res. 49, 409–436.

Hicks, G., Pratt, R.G., 2001. Reflection waveform inversion using local descent methods: estimating attenuation and velocity over a gas-sand deposit. Geophysics 66, 598–612.

Hildebrand, J.A., Dorman, L.M., Hammer, P.T.C., Schreiner, A.E., Cornuelle, B.D., 1989. Seismic tomography of Jasper Seamount. Geophys. Res. Lett. 16, 1355–1358.

Hirahara, K., 1988. Detection of three-dimensional velocity anisotropy. Phys. Earth Planet. Inter. 51, 71–85.

Hole, J.A., 1992. Nonlinear high-resolution three-dimensional travel-time tomography. J. Geophys. Res. 97, 6553–6562.

Hole, J.A., Zelt, B.C., 1995. 3-D finite-difference reflection travel times. Geophys. J. Int. 121, 427–434.

Humphreys, E., Clayton, R.W., 1988. Adaption of back projection tomography to seismic travel time problems. J. Geophys. Res. 93, 1073–1085.

Humphreys, E.D., Clayton, R.W., 1990. Tomographic image of the Southern California Mantle. J. Geophys. Res. 95, 19725–19746.

Hung, S.H., Dahlen, F.A., Nolet, G., 2000. Fréchet kernels for finite-frequency travel-times. II. Examples. Geophys. J. Int. 141, 175–203.

Hung, S.H., Shen, Y., Chiao, L.Y., 2004. Imaging seismic velocity structure beneath the Iceland hotspot: a finite frequency approach. J. Geophys. Res. 109, B08305.

Ishii, M., Tromp, J., 2004. Constraining large-scale mantle heterogeneity using mantle and inner-core sensitive normal modes. Phys. Earth Planet. Inter. 146, 113–124.

Iyer, H., Hirahara, K., 1993. Seismic Tomography: Theory and Practice. Chapman & Hall, London.

Jackson, I., Paterson, M., Fitz Gerald, J., 1992. Seismic wave dispersion and attenuation in Aheim dunite: an experimental study. Geophys. J. Int. 108, 517–534.

Jaiswal, P., Zelt, C.A., Bally, A.W., Dasgupta, R., 2008. 2-D traveltime and waveform inversion for improved seismic imaging: Naga Thrust and Fold Belt, India. Geophys. J. Int. 173, 642–658.

Jiang, G.S., Peng, D.P., 2000. Weighted ENO schemes for Hamilton–Jacobi equations. SIAM J. Sci. Comput. 21, 2126–2143.

Jiang, G.S., Shu, C., 1996. Efficient implementation of weighted ENO schemes. J. Comp. Phys. 126, 202–228.

Julia, J., Ammon, C.J., Hermann, R.B., 2003. Lithospheric structure of the Arabian Shield from the joint inversion of receiver functions and surface-wave group velocities. Tectonophysics 371, 1–21.

Julian, B.R., Gubbins, D., 1977. Three-dimensional seismic ray tracing. J. Geophys. 43, 95–113.

Kanasewich, E., Burianyk, M.J.A., Ellis, R.M., Clowes, R.M., White, D.J., Lôté, T., Forsyth, D.A., Luetgert, J.A., Spence, G.D., 1994. Crustal velocity structure of the Omineca Belt, southwestern Canadian Cordillera. J. Geophys. Res. 99, 2653–2670.

Kanasewich, E.R., Chiu, S.K.L., 1985. Least-squares inversion of spatial seismic refraction data. Bull. Seismol. Soc. Am. 75, 865–880.

Kang, T.-S., Shin, J.S., 2006. Surface-wave tomography from ambient seismic noise of accelerograph networks in southern Korea. Geophys. Res. Lett. 33, doi:10.1029/2006GL027044.

Karason, H., van der Hilst, R.D., 2001. Improving global tomography models of P-wavespeed. I. Incorporation of differential travel times for refracted and diffracted core phases (PKP, Pdiff). J. Geophys. Res. 106, 6569–6587.

Kelley, K., Plank, T., submitted for publication. Geochemistry of the Mariana Arc System. Geochem. Geophys. Geosyst.

Kennett, B.L.N., Engdahl, E.R., Buland, R., 1995. Constraints on seismic velocities in the earth from travel times. Geophys. J. Int. 122, 108–124.

Kennett, B.L.N., 1998. Seismic Wave Propagation and Seismic Tomography. Research School of Earth Sciences. Institute of Advanced Studies, The Australian National University, Canberra.

Kennett, B.L.N., Sambridge, M.S., Williamson, P.R., 1988. Subspace methods for large scale inverse problems involving multiple parameter classes. Geophys. J. 94, 237–247.

Khan, A., Connolly, J.A.D., Taylor, S.R., 2008. Inversion of seismic and geodetic data for the major element chemistry and temperature of the Earth's mantle. J. Geophys. Res. 113, doi:10.1029/2007JB005239.

Kim, S., Cook, R., 1999. 3D traveltime computation using second-order ENO scheme. Geophysics 64, 1867–1876.

Klimeš, L., Kvasnička, M., 1994. 3-D network ray tracing. Geophys. J. Int. 116, 726–738.

Koketsu, K., Sekine, S., 1998. Pseudo-bending method for three-dimensional seismic ray tracing in a spherical earth with discontinuities. Geophys. J. Int. 132, 339–346.

Komatitsch, D., Ritsema, J., Tromp, J., 2002. The spectral-element method, beowulf computing, and global seismology. Science 298, 1737–1742.

Komatitsch, D., Tromp, J., 1999. Introduction to the spectral-element method for 3-D seismic wave propagation. Geophys. J. Int. 139, 806–822.

Komatitsch, D., Tsuboi, S., Tromp, J., 2005. The spectral-element method in seismology. In: Levander, A., Nolet, G. (Eds.), Seismic Earth: Array Analysis of Broadband Seismograms, vol. 157. American Geophysical Union, pp. 205–227.

Komatitsch, D., Vilotte, J.P., 1998. The spectral element method: an efficient tool to simulate the seismic response of 2D and 3D geological structured. Bull. Seismol. Soc. Am. 88, 368–392.

Korenaga, J., Holbrook, W.S., Kent, G.M., Kelemen, R.S., Detrick, R.S., Larsen, H.C., Hopper, J.R., Dahl-Jensen, T., 2000. Crustal structure of the southeast Greenland margin from joint refraction and reflection tomography. J. Geophys. Res. 105, 21,591–21,614.

Kustowski, B., Ekstrom, G., Dziewonski, A.M., 2008. Anisotropic shear-wave velocity structure of the Earths mantle: a global model. J. Geophys. Res. 113, B06306, doi:10.1029/2007JB005169.

Lambaré, G., Lucio, P.S., Hanyga, A., 1996. Two-dimensional multivalued traveltime and amplitude maps by uniform sampling of a ray field. Geophys. J. Int. 125, 584–598.

Langan, R.T., Lerche, I., Cutler, R.T., 1985. Tracing of rays through heterogeneous media: an accurate and efficient procedure. Geophysics 50, 1456–1465.

Laske, G., Masters, G., 1996. Constraints on global phase velocity maps from long-period polarization data. J. Geophys. Res. 101, 16059–16075.

Lawson, C.L., Hanson, R.J., 1974. Solving Least Squares Problems. Prentice-Hall, New Jersey.

Lebedev, S., Boonen, J., Trampert, J., 2009. Seismic structure of precam-brian lithosphere: new constraints from broad-band surface-wave dispersion. Lithos 109, 96–111.

Lebedev, S., van der Hilst, R.D., 2008. Global upper-mantle tomography with the automated multimode invesion of surface and S-wave forms. Geo-phys. J. Int. 173, 505–518.

Lees, J.M., Crosson, R.S., 1989. Tomographic inversion for three-dimensional velocity structure at Mount St. Helens using earthquake data. J. Geophys. Res. 94, 5716–5728.

Lees, J.M., VanDecar, J.C., 1991. Seismic tomography constrained by Bouguer gravity anomalies: applications in western Washington. Pageoph 135, 31–52.

Lei, J., Zhao, D., 2007. Teleseismic P-wave tomography and the upper mantle structure of the central Tien Shan orogenic belt. Phys. Earth Planet. Inter. 162, 165–185.

Lévêque, J.J., Rivera, L., Wittlinger, G., 1993. On the use of the checkerboard test to assess the resolution of tomographic inversions. Geophys. J. Int. 115, 313–318.

Li, A., Burke, K., 2006. Upper mantle structure of southern Africa from Rayleigh Wave Tomography. J. Geophys. Res. 111, doi:10.1029/2006JB004321.

Li, X.D., Giardini, D., Woodhouse, J., 1991. Large-scale three-dimensional even-degree structure of the earth from splitting of long-period normal modes. J. Geophys. Res. 96, 551577.

Li, X.G., Romanowicz, B., 1995. Comparison of global waveform inversions with and without considering cross-branch modal coupling. Geophys. J. Int. 121, 695–709.

Lippitsch, R., Kissling, E., Ansorge, J., 2003. Upper mantle structure beneath the Alpine orogen from high-resolution teleseismic tomography. J. Geophys. Res. 108, 2376, doi:10.1029/2002JB002016.

Liu, Q., Tromp, J., 2008. Finite-frequency sensitivity kernels for global seismic wave propagation based upon adjoint methods. Geophys. J. Int. 174, 265–286.

Liu, X., Osher, S., Chan, T., 1994. Weighted essentially non-oscillatory schemes. J. Comp. Phys. 115, 200–212.

Lobkis, O.I., Weaver, R.L., 2001. On the emergence of the Green's function in the correlations of a diffuse field. J. Acoust. Soc. Am. 110, 3011–3017.

Long, M.D., de Hoop, M.V., Van Der Hilst, R.D., 2008. Wave-equation shear wave splitting tomography. Geophys. J. Int. 172, 311–330.

Loris, I., Nolet, G., Daubechies, I., Dahlen, F., 2007. Tomographic inversion using l1-norm regularization of wavelet coefficients. Geophys. J. Int. 170, 359–370.

Louden, K.E., Fan, J., 1998. Crustal structures of Grenville, Makkovik, and southern Nain provinces along the Lithoprobe ECSOOT Transect: regional seismic refraction and gravity models and their tectonic implications. Can. J. Earth Sci. 35, 583–601.

Lukas, M.A., 2008. Strong robust generalized cross-validation for choosing the regularization parameter. Inverse Probl. 24, 034006.

Luo, Y., Shuster, G.T., 1991. Wave-equation traveltime inversion. Geophysics 56, 645–653.

Lutter, W.J., Nowack, R.L., 1990. Inversion for crustal structure using reflections from the PASSCAL Ouachita experiment. J. Geophys. Res. 95, 4633–4646.

Madge, L.S., Sparks, D.W., 1997. Three-dimensional mantle upwelling, melt generation, and melt migration beneath segment slow spreading ridges. J. Geophys. Res. 102, 20571–20583.

Marfurt, K.J., 1984. Accuracy of finite-difference and finite-element modeling of the scalar and elastic wave equations. Geophysics 49, 533–549.

Marquering, H., Dahlen, F.A., Nolet, G., 1999. Three-dimensional sensitivity kernels for finite-frequency travel times: the banana–doughnut paradox. Geophys. J. Int. 137, 805–815.

Marquering, H., Nolet, G., Dahlen, F.A., 1998. Three-dimensional waveform sensitivity kernels. Geophys. J. Int. 132, 521–534.

Martin, M., Ritter, J.R.R., the CALIXTO working group, 2005. High-resolution tele-seismic body-wave tomography beneath SE Romania. I. Implications for the three-dimensional versus one-dimensional crustal correction strategies with a new crustal velocity model. Geophys. J. Int. 162, 448–460.

McCaughey, M., Singh, S.C., 1997. Simultaneous velocity and interface tomography of normal-incidence and wide-aperture seismic traveltime data. Geophys. J. Int. 131, 87–99.

McMechan, G.A., 1983. Seismic tomography in boreholes. Geophys. J. Royal Astr. Soc. 74, 601–612.

McMechan, G.A., 1987. Cross-hole tomography for strongly variable media with applications to scale model data. Bull. Seismol. Soc. Am. 77, 1945–1960.

McQueen, H.W.S., Lambeck, K., 1996. Determination of crustal structure in central Australia by inversion of traveltime residuals. Geophys. J. Int. 126, 645–662.

Menke, W., 1989. Geophysical Data Analysis: Discrete Inverse Theory. Academic Press, New York.

Michelini, A., 1995. An adaptive-grid formalism for traveltime tomography. Geophys. J. Int. 121, 489–510.

Mjelde, R., Digranes, P., Shimamura, H., Shiobara, H., Kodaira, S., Brekke, H., Egebjerg, T., Sørenes, N., Thorbjørnsen, S., 1998. Crustal structure of the northern part of the Vøring Basin, mid-Norway margin, from wide-angle seismic and gravity data. Tectonophysics 293, 175–205.

Montagner, J.-P., Nataf, H.-C., 1986. A simple method for inverting the aziumthal anisotropy of surface waves. J. Geophys. Res. 91, 511–520.

Montagner, J.P., Tanimoto, T., 1990. Global anisotropy in the upper mantle inferred from the regionalization of phase velocities. J. Geophys. Res. 95, 4797–4819.

Montagner, J.P., Tanimoto, T., 1991. Global upper mantle tomography of seismic velocities and anisotropies. J. Geophys. Res. 96, 20337–20351.

Monteiller, V., Got, J.L., 2005. An efficient algorithm for double-difference tomography and location in heterogeneous media, with an application to the Kilauea volcano. J. Geophys. Res. 110, doi:10.1029/2004JB003466.

Montelli, R., Nolet, G., Dahlen, F.A., 2006. Comment on 'Banana–doughnut kernels and mantle tomography' by van der Hilst and de Hoop. Geophys. J. Int. 167, 1204–1210.

Montelli, R., Nolet, G., Dahlen, F.A., Masters, G., Engdahl, E.R., Hung, S.H., 2004. Finite-frequency tomography reveals a variety of plumes in the mantle. Science 303, 338–343.

Moore, G.E., 1965. Cramming more components onto integrated circuits. Electronics 38, 114–117.

Morgan, R.P.L., Barton, P.J., Warner, M., Morgan, J., Price, C., Jones, K., 2000. Lithospheric structure north of Scotland. I. P-wave modelling, deep reflection profiles and gravity. Geophys. J. Int. 142, 716–736.

Morozov, I.B., Smithson, S.B., Hollister, L.S., Diebold, J.B., 1998. Wide-angle seismic imaging across accreted terranes, southeastern Alaska and western British Columbia. Tectonophysics 299, 281–296.

Mosegaard, K., Sambridge, M., 2002. Monte Carlo analysis of inverse problems. Inverse Probl. 18, R29–R54.

Moser, T.J., 1991. Shortest path calculation of seismic rays. Geophysics 56, 59–67.

Nakanishi, I., 1985. Three-dimensional structure beneath the Hokkaido-Tohoku region as derived from a tomographic inversion of P-arrival times. J. Phys. Earth 33, 241–256.

Nakanishi, I., Yamaguchi, K., 1986. A numerical experiment on nonlinear image reconstruction from first-arrival times for two-dimensional island arc structure. J. Phys. Earth 34, 195–201.

Nataf, H.-C., Nakanishi, I., Anderson, D.L., 1984. Anisotropy and shear-velocity heterogeneities in the upper mantle. Geophys. Res. Lett. 11, 109–112.

Neele, F., VanDecar, J., Snieder, R., 1993. The use of P wave amplitude data in a joint inversion with travel times for upper mantle velocity structure. J. Geophys. Res. 98, 12033–12054.

Nissen-Meyer, T., Dahlen, F.A., Fournier, A., 2007. Spherical-earth Fréchet sensitivity kernels. Geophys. J. Int. 168, 1051–1066.

Nolet, G., 1985. Solving or resolving inadequate and noisy tomographic systems. J. Comp. Phys. 61, 463–482.

Nolet, G., 1987. Waveform tomography. In: Nolet, G. (Ed.), Seismic Tomography: With Applications in Global Seismology and Exploration Geophysics. D. Reidel, Dordrecht, pp. 301–322.

Nolet, G., 1990. Partitioned waveform inversion and two-dimensional structure under the network of autonomously recording seismographs. J. Geo-phys. Res. 95, 8499–8512.

Nolet, G., 2008. A Breviary of Seismic Tomography: Imaging the Interior of the Earth and the Sun. Cambridge University Press, Cambridge.

Nolet, G., Montelli, R., 2005. Optimal parameterization of tomographic models. Geophys. J. Int. 161, 1–8.

Nolet, G., Montelli, R., Virieux, J., 1999. Explicit, approximate expressions for the resolution and a posteriori covariance of massive tomographic systems. Geophys. J. Int. 138, 36–44.

Oncescu, M.C., Burlacu, V., Anghel, M., Smalbergher, V., 1984. Three-dimensional P-wave velocity image under the Carpathian Arc. Tectono-physics 106, 305–319.

Osher, S., Cheng, L.-T., Kang, M., Shim, H., Tsai, Y.-H., 2002. Geometric optics in a phase-space-based level set and Eulerian framework. J. Comp. Phys. 179, 622–648.

Özalaybey, S., Savage, M.K., Sheehan, A.F., Louie, J.N., Brune, J.N., 1997. Shear-wave velocity structure in the northern Basin and Range province from the combined analysis of receiver functions and surface waves. Bull. Seismol. Soc. Am. 87, 183–189.

Panning, M., Romanowicz, B., 2006. A three-dimensional radially anisotropic model of shear velocity in the whole mantle. Geophys. J. Int. 167, 361–379.

Panning, M.P., Capdeville, Y., Romanowicz, B.A., 2009. Seismic waveform modelling in a 3-D Earth using the Born approximation: potential shortcomings and a remedy. Geophys. J. Int. 177, 161–178.

Parmentier, E., Phipps Morgan, J., 1990. Spreading rate dependence of three-dimensional structure in oceanic spreading centers. Nature 348, 325–328.

Parsons, T., Zoback, M.L., 1997. Three-dimensional upper crustal velocity structure beneath San Francisco Peninsula. Calif. J. Geophys. Res. 102, 5473–5490.

Pasyanos, M.E., Nyblade, A.A., 2007. A top to bottom lithospheric study of Africa and Arabia. Tectonophysics 444, 27–44.

Pearce, J.A., Stern, R.J., Bloomer, S., Fryer, P., 2005. Geochemical mapping of the Mariana arc-basin system: implications for the nature and distribution of subduction components. Geochem. Geophy. Geosyst. 6, Q07006, doi:10.1029/2004GC000895.

Pedersen, H.A., Coutant, O., Deschamps, A., Soulage, M., Cotte, N., 2003. Measuring surface wave phase velocities beneath small broadband arrays: test of an improved algorithm and application to the french alps. Geophys. J. Int. 154, 903–912.

Pedersen, H.A., Fishwick, S., Snyder, D.B., 2009. A comparison of cratonic roots through consistent analysis of seismic surface waves. Lithos 109, 81–95.

Pereyra, V., 1996. Modelling, ray tracing, and block nonlinear travel-time inversion in 3D. Pure Appl. Geophys. 148, 345–386.

Pereyra, V., Lee, W.H.K., Keller, H.B., 1980. Solving two-point seismic-ray tracing problems in a heterogeneous medium. Bull. Seismol. Soc. Am. 70, 79–99.

Plomerová, J., Babuška, V., Kozlovskaya, E., Vecsey, L., Hyvönen, 2008. Seismic anisotropy: a key to resolve fabrics of mantle lithosphere of Fennoscandia. Tectonophysics 462, 125–136.

Plomerová, J., Babuška, V., Vecsey, L., Kouba, D., TOR Working Group, 2002. Seismic anisotropy of the lithosphere around the Trans-European Suture Zone (TESZ) based on teleseismic body-wave data of the TOR experiment. Tectonophysics 360, 89–114.

Podvin, P., Lecomte, I., 1991. Finite difference computation of traveltimes in very contrasted velocity models: a massively parallel approach and its associated tools. Geophys. J. Int. 105, 271–284.

Popovici, A.M., Sethian, J.A., 2002. 3-D imaging using higher order fast marching traveltimes. Geophysics 67, 604–609.

Pozgay, S.H., Wiens, D.A., Conder, J., Shiobara, H., Sugioka, H., 2009. Seismic attenuation tomography of the Mariana subduction system: implications for thermal structure, volatile distribution, and slow spreading dynamics. Geochem. Geophys. Geosyst. 10, doi:10.1029/2008GC002313.

Pratt, R.G., 1990. Frequency-domain elastic wave modelling by finite differences: a tool for crosshole seismic imaging. Geophysics 55, 626–632.

Pratt, R.G., 1999. Seismic waveform inversion in the frequency domain. Part 1. Theory and verification in a physical scale model. Geophysics 64, 888–901.

Pratt, R.G., Goulty, N.R., 1991. Combining wave-equation imaging with traveltime tomography to form high-resolution images from crosshole data. Geophysics 56, 208–224.

Pratt, R.G., Shipp, R.M., 1999. Seismic waveform inversion in the frequency domain. Part 2. Fault delineation in sediments using crosshole data. Geophysics 64, 902–914.

Pratt, R.G., Song, Z.M., Williamson, P., Warner, M., 1996. Two-dimensional velocity models from wide-angle seismic data by wavefield inversion. Geophys. J. Int. 124, 323–340.

Pratt, R.G., Worthington, M.H., 1988. The application of diffraction tomography to cross-hole seismic data. Geophysics 53, 1284–1294.

Pratt, R.G., Worthington, M.H., 1990. Inverse theory applied to multi-source cross-hole tomography. Part 1. Acoustic wave equation method. Geophys. Prospect. 38, 311–330.

Priestley, K., McKenzie, D., Debayle, E., Pilidou, S., 2008. The African upper mantle and its relationship to tectonics and surface geology. Geophys. J. Int. 175, 1108–1126.

Pullammanappallil, S.K., Louie, J.N., 1993. Inversion of seismic reflection traveltimes using a nonlinear optimization scheme. Geophysics 58, 1607–1620.

Pulliam, J., Snieder, R., 1996. Fast, efficient calculation of rays and travel times with ray perturbation theory. J. Acoust. Soc. Am. 99, 383–391.

Qian, J., Symes, W.W., 2002. An adaptive finite-difference method for traveltimes and amplitudes. Geophysics 67, 167–176.

Qin, F., Luo, Y., Olsen, K.B., Cai, W., Schuster, G.T., 1992. Finite-difference solution of the eikonal equation along expanding wavefronts. Geophysics 57, 478–487.

Rawlinson, N., Hauser, J., Sambridge, M., 2007. Seismic ray tracing and wavefront tracking in laterally heterogeneous media. AG, 203–267.

Rawlinson, N., Houseman, G.A., 1998. Inversion for interface structure using teleseismic traveltime residuals. Geophys. J. Int. 133, 756–772.

Rawlinson, N., Houseman, G.A., Collins, C.D.N., 2001a. Inversion of seismic refraction and wide-angle reflection traveltimes for 3-D layered crustal structure. Geophys. J. Int. 145, 381–401.

Rawlinson, N., Houseman, G.A., Collins, C.D.N., Drummond, B.J., 2001b. New evidence of Tasmania's tectonic history from a novel seismic experiment. Geophys. Res. Lett. 28, 3337–3340.

Rawlinson, N., Kennett, B.L.N., 2004. Rapid estimation of relative and absolute delay times across a network by adaptive stacking. Geophys. J. Int. 157, 332–340.

Rawlinson, N., Kennett, B.L.N., 2008. Teleseismic tomography of the upper mantle beneath the southern Lachan Orogen. Aust. Phys. Earth Planet. Inter. 167, 84–97.

Rawlinson, N., Kennett, B.L.N., Heintz, M., 2006a. Insights into the structure of the upper mantle beneath the Murray Basin from 3D teleseismic tomography. Aust. J. Earth Sci. 53, 595–604.

Rawlinson, N., Reading, A.M., Kennett, B.L.N., 2006b. Lithospheric structure of Tasmania from a novel form of teleseismic tomography. J. Geophys. Res. 111, doi:10.1029/2005JB003803.

Rawlinson, N., Sambridge, M., 2003a. Irregular interface parameterization in 3-D wide-angle seismic traveltime tomography. Geophys. J. Int. 155, 79–92.

Rawlinson, N., Sambridge, M., 2003b. Seismic traveltime tomography of the crust and lithosphere. Adv. Geophys. 46, 81–198.

Rawlinson, N., Sambridge, M., 2004a. Multiple reflection and transmission phases in complex layered media using a multistage fast marching method. Geophysics 69, 1338–1350.

Rawlinson, N., Sambridge, M., 2004b. Wavefront evolution in strongly heterogeneous layered media using the fast marching method. Geophys. J. Int. 156, 631–647.

Rawlinson, N., Sambridge, M., Saygin, E., 2008. A dynamic objective function technique for generating multiple solution models in seismic tomography. Geophys. J. Int. 174, 295–308.

Rawlinson, N., Urvoy, M., 2006. Simultaneous inversion of active and passive source datasets for 3-D seismic structure with application to Tasmania. Geophys. Res. Lett. 33, doi:10.1029/2006GL028105.

Reed, A.R., 2001. Pre-Tabberabberan deformation in eastern Tasmania: a southern extension of the Benambran Orogeny. Aust. J. Earth Sci. 48, 785–796.

Reid, F.J.L., Woodhouse, J.H., van Heist, H.H., 2001. Upper mantle attenuation and velocity sstructure from measurements of differential S phases. Geophys. J. Int. 145, 615–630.

Resovsky, J.S., Ritzwoller, M.H., 1999. A degree 8 mantle shear velocity model from normal mode observations below 3 mHz. J. Geophys. Res. 104, 9931014.

Riahi, M.A., Juhlin, C., 1994. 3-D interpretation of reflected arrival times by finite-difference techniques. Geophysics 59, 844–849.

Riahi, M.A., Lund, C.E., Pederson, L.B., 1997. Three-dimensional image of the Moho undulations beneath the Gulf of Bothnia using wide-angle seismic data. Geophys. J. Int. 129, 461–471.

Ritsema, J., Nyblade, A.A., Owens, T.J., Langston, C.A., VanDecar, J.C., 1998. Upper mantle seismic velocity structure beneath Tanzania, east Africa: implications for the stability of cratonic lithosphere. J. Geophys. Res. 103, 21201–21213.

Ritsema, J., van Heijst, H.-J., Woodhouse, J.H., 1999. Complex shear wave velocity structure imaged beneath Africa and Iceland. Science 286, 1925–1928.

Ritsema, J., van Heijst, H.J., Woodhouse, J.H., 2004. Global transition zone tomography. J. Geophys. Res. 109, doi:10.1029/2003JB002610.

Ritzwoller, M.H., Levshin, A.L., 1998. Eurasian surface wave tomography: group velocities. J. Geophys. Res. 103, 4839–4878.

Romanowicz, B., 1995. A global tomographic model of shear attenuation in the upper mantle. J. Geophys. Res. 100, 12375–12394.

Romanowicz, B., 2003. Global mantle tomography: progress status in the past 10 years. Ann. Rev. Earth Planet. Sci. 31, 303–328.

Romanowicz, B., Gung, Y., 2002. Superplumes from the core–mantle boundary to the lithosphere: implications for heat flux. Science 296, 513–516.

Ronchi, C., Iacono, R., Paolucci, P.S., 1996. The "Cubed Sphere": a new method for the solution of partial differential equations in spherical geometry. J. Comp. Phys. 124, 93–114.

Roth, E.G., Wiens, D.A., Dorman, L.M., Hildebrand, J., Webb, S.C., 1999. Seismic attenuation tomography of the Tonga-Fiji region using phase pair methods. J. Geophys. Res. 104, 4795–4809.

Roy, L., Sen, M.K., McIntosh, K., Stoffa, P.L., Nakamura, Y., 2005. Joint inversion of first arrival seismic travel-time and gravity data. J. Geophys. Eng. 2, 277–289.

Sabra, K.G., Gerstoft, P., Roux, P., Kuperman, W.A., Fehler, M.C., 2005. Surface wave tomography from microseisms in Southern California. Geophys. Res. Lett. 32, doi:10.1029/2005GL023155.

Sadeghi, H., Suzuki, S., Takenaka, H., 1999. A two-point, three-dimensional seismic ray tracing using genetic algorithms. Phys. Earth Planet. Inter. 113, 355–365.

Saltzer, R.L., Humphreys, E.D., 1997. Upper mantle P wave velocity structure of the eastern Snake River Plain and its relationship to geodynamic models of the region. J. Geophys. Res. 102, 11829–11841.

Sambridge, M., Faletic, R., 2003. Adaptive whole Earth tomography. Geochem. Geophys. Geosyst. 4, doi:10.1029/2001GC000213.

Sambridge, M., Gudmundsson, O., 1998. Tomographic systems of equations with irregular cells. J. Geophys. Res. 103, 773–781.

Sambridge, M., Mosegaard, K., 2001. Monte Carlo methods in geophysical inverse problems. Rev. Geophys. 40, doi:10.1029/2000RG000089.

Sambridge, M., Rawlinson, N., 2005. Seismic tomography with irregular meshes. In: Levander, A., Nolet, G. (Eds.), Seismic Earth: Array Analysis of Broadband Seismograms, vol. 157. American Geophysical Union, pp. 49–65.

Sambridge, M., Rickwood, P., Rawlinson, N., Sommacal, S., 2008. Automatic differentiation in geophysical inverse problems. Geophys. J. Int. 170, 1–8.

Sambridge, M.S., 1990. Non-linear arrival time inversion: constraining velocity anomalies by seeking smooth models in 3-D. Geophys. J. Int. 102, 653–677.

Sambridge, M.S., Braun, J., McQueen, H., 1995. Geophysical parametriza-tion and interpolation of irregular data using natural neighbours. Geophys. J. Int. 122, 837–857.

Sanders, C.O., 1993. Local earthquake tomography: attenuation-theory and results. In: Iyer, H.M., Hirahara, K. (Eds.), Seismic Tomography: Theory and Practice. Chapman & Hall, London, pp. 676–694.

Sato, T., Kosuga, M., Tanaka, K., 1996. Tomographic inversion for P wave velocity structure beneath the northeastern Japan arc using local and tele-seismic data. J. Geophys. Res. 101, 17597–17615.

Saygin, E., Kennett, B., 2009. Ambient seismic noise tomography of Australian continent. Tectonophysics, doi:10.1016/j.tecto.2008.11.013.

Scales, J.A., 1987. Tomographic inversion via the conjugate gradient method. Geophysics 52, 179–185.

Scales, J.A., Snieder, R., 1997. To Bayes or not to Bayes. Geophysics 62, 1045–1046.

Schefter, J., 1979. Supercomputer: incredible Cray-1 cruises at 80 milliono operations a second. Popular Sci. (June), 86–89.

Scott, J.S., Masters, T.G., Vernon, F.L., 1994. 3-D velocity structure of the San Jacinto fault zone near Anza, California-I. P waves. Geophys. J. Int. 119, 611–626.

Sebai, A., amd Stutzmann, E., Montagner, J.P., Sicilia, D., Beucler, E., 2006. Anisotropic structure of the African upper mantle from Rayleigh and Love wave tomography. Phys. Earth Planet. Inter. 155, 48–62.

Selby, N.D., Woodhouse, J.H., 2002. The Q structure of the upper mantle: constraints from Rayleigh wave amplitudes. J. Geophys. Res. 107, 933–940.

Sethian, J.A., 1996. A fast marching level set method for monotonically advancing fronts. Proc. Natl. Acad. Sci. 93, 1591–1595.

Sethian, J.A., Popovici, A.M., 1999. 3-D traveltime computation using the fast marching method. Geophysics 64, 516–523.

Shapiro, N.M., Campillo, M., 2004. Emergence of broadband Rayleigh waves from correlations of the ambient seismic noise. Geophys. Res. Lett. 31, doi:10.1029/2004GL019491.

Shapiro, N.M., Campillo, M., Stehly, L., Ritzwoller, M.H., 2005. High-resolution surface wave tomography from ambient seismic noise. Science 307, 1615–1618.

Shapiro, N.M., Ritzwoller, M.H., 2002. Monte-Carlo inversion for a global shear-velocity model of the crust and upper mantle. Geophys. J. Int. 151, 88–105.

Shito, A., Karato, S.-I., Matsukage, K.N., Nishibara, Y., 2006. Towards mapping the three-dimensional distribution of water in the upper mantle from velocity and attenuation tomography. In: Earth's Deep Water Cycle. Geophysical Monograph Series, vol. 168. American Geophysical Union, Washington, DC, pp. 225–236.

Shomali, Z.H., Roberts, R.G., the TOR Working Group, 2002. Non-linear body wave teleseismic tomography along the TOR array. Geophys. J. Int. 148, 562–574.

Shu, C.-W., Osher, S., 1988. Efficent implementation of essentially non-oscillatory shock-capturing schemes. J. Comp. Phys. 77, 439–471.

Shu, C.-W., Osher, S., 1989. Efficent implementation of essentially non-oscillatory shock-capturing schemes, II. J. Comp. Phys. 83, 32–78.

Sigloch, K., McQuarrie, N., Nolet, G., 2008. Two-stage subduction history under North America inferred from multiple-frequency tomography. Nature Geosci. 1, 458–462.

Simons, F., Zielhuis, A., van der Hilst, R.D., 1999. The deep structure of the Australian continent from surface wave tomography. Lithos 48, 17–43.

Simons, F.J., van der Hilst, R.D., Montagner, J.-P., Zielhuis, A., 2002. Mul-timode Rayleigh wave inversion for heterogeneity and azimuthal anisotropy of the Australian upper mantle. Geophys. J. Int. 151, 738–754.

Simons, F.J., Nolet, G., Gergief, P., Babcock, J.M., Regier, L.A., Davis, R.E., 2009. On the potential of recording earthquakes for global seismic tomography by low-cost autonomous instruments in the oceans. J. Geophys. Res. 114, B05307, doi:10.1029/2008JB006088.

Sirgue, L., Pratt, R.G., 2004. Efficient waveform inversion and imaging: a strategy for selecting temporal frequencies. Geophysics 69, 231–248.

Slawinski, M.A., Slawinski, R.A., Brown, R.J., Parkin, J.M., 2000. A generalized form of snell's law in anisotropic media. Geophysics 65, 632–637.

Smith, W.H.F., Wessel, P., 1990. Gridding with continuous curvature splines in tension. Geophysics 55, 293–305.

Snieder, R., 1988a. Large-scale waveform inversions of surface waves for lateral heterogeneity. 1. Theory and numerical examples. J. Geophys. Res. 93, 12055–12065.

Snieder, R., 1988b. Large-scale waveform inversions of surface waves for lateral heterogeneity. 2. Application to surface waves in Europe and the Mediterranean. J. Geophys. Res. 93, 12067–12080.

Snieder, R., 1993. Global inversions using normal modes and long-period surface waves. In: Iyer, H.M., Hirahara, K. (Eds.), Seismic Tomography: Theory and Practice. Chapman & Hall, London, pp. 23–63.

Snieder, R., 2004. Extracting the Green's function from the corelation of coda waves: a derivation based on stationary phase. Phys. Rev. E 69, doi:10.1103/PhysRevE.69.046610.

Snieder, R., Sambridge, M., 1992. Ray perturbation theory for traveltimes and ray paths in 3-D heterogeneous media. Geophys. J. Int. 109, 294–322.

Song, Z.M., Williamson, P.R., Pratt, R.G., 1995. Frequency-domain acoustic-wave modelling and inversion of cross-hole data. Part II. Inversion method, synthetic experiments and real-data results. Geophysics 60, 796–809.

Spakman, W., Bijwaard, H., 2001. Optimization of cell parameterizations for tomographic inverse problems. Pure Appl. Geophys. 158, 1401–1423.

Staples, R.K., White, R.S., Brandsdóttir, Menke, W., Maguire, P.K.H., McBride, J.H., 1997. Färoe-Iceland Ridge experiment. 1. Crustal structure of northeastern Iceland. J. Geophys. Res. 102, 7849–7866.

Steck, L.K., Thurber, C.H., Fehler, M., Lutter, W.J., Roberts, P.M., Baldridge, W.S., Stafford, D.G., Sessions, R., 1998. Crust and upper mantle *P* wave velocity structure beneath Valles caldera, New Mexico: results from the Jemez teleseismic tomography experiment. J. Geophys. Res. 103, 24301–24320.

Steinhoff, J., Fan, M., Wang, L., 2000. A new Eulerian method for the computation of propagating short acoustic and electromagnetic pulses. J. Comp. Phys. 157, 683–706.

Štekl, I., Pratt, R.G., 1998. Accurate visco-elastic modeling by frequency-domain finite differences using rotated operators. Geophysics 63, 1779–1794.

Su, W.-J., Dziewonski, A.M., 1997. Simultaneous inversion for 3-D variations in shear and bulk velocity in the mantle. Phys. Earth Planet. Inter. 100, 135–156.

Symes, W.W., Qian, J., 2003. A slowness matching eulerian method for multivalued solutions of eikonal equaions. SIAM J. Sci. Comput. 19, 501–526.

Tanimoto, T., Anderson, D.L., 1984. Mapping convection in the mantle. Geophys. Res. Lett. 11, 287–290.

Tanimoto, T., Anderson, D.L., 1985. Lateral heterogeneity and azimuthal anisotropy of the upper mantle—Love and Rayleigh waves 100–250 sec. J. Geophys. Res. 90, 1842–1858.

Tarantola, A., 1984. Inversion of seismic reflection data in the acoustic approximation. Geophysics 49, 1259–1266.

Tarantola, A., 1987. Inverse Problem Theory. Elsevier, Amsterdam.

Tarantola, A., Nercessian, A., 1984. Three-dimensional inversion without blocks. Geophys. J. Royal Astr. Soc. 76, 299–306.

Thomson, C.J., Gubbins, D., 1982. Three-dimensional lithospheric modelling at NOR-SAR: linearity of the method and amplitude variations from the anomalies. Geophys. J. Royal Astr. Soc. 71, 1–36.

Thurber, C.H., 1983. Earthquake locations and three-dimensional crustal structure in the Coyote Lake area, central California. J. Geophys. Res. 88, 8226–8236.

Thurber, C.H., Ellsworth, W.L., 1980. Rapid solution of ray tracing problems in heterogeneous media. Bull. Seismol. Soc. Am. 70, 1137–1148.

Tian, Y., Hung, S.-H., Nolet, G., Montelli, R., Dahlen, F.A., 2007a. Dynamic ray tracing and traveltime corrections for global seismic tomography. J. Comp. Phys. 226, 672–687.

Tian, Y., Montelli, R., Nolet, G., Dahlen, F.A., 2007b. Computing traveltime and amplitude sensitivity kernels in finite frequency tomography. J. Comp. Phys. 226, 2271–2288.

Tikhotsky, S., Achauer, U., 2008. Inversion of controlled-source seismic tomography and gravity data with self-adaptive wavelet parameterization of velocities and interfaces. Geophys. J. Int. 172, 619–630.

Tkalčić, H., Pasyanos, M., Rodgers, A., Gok, W., Walter, W., Al-amri, A., 2006. A multi-step approach in joint modeling of surface wave dispersion and teleseismic receiver functions: implications for lithospheric structure of the Arabian peninsula. J. Geophys. Res. 111, doi:10.1029/2005JB004130.

Toomey, D.R., Solomon, S.C., Purdy, G.M., 1994. Tomographic imaging of the shallow crustal structure of the East Pacific Rise at 9°30′N. J. Geophys. Res. 99, 24,135–24,157.

Trampert, J., Spetzler, J., 2006. Surface wave tomography: finite-frequency effects lost in the null space. Geophys. J. Int. 164, 394–400.

Trampert, J., Woodhouse, J.H., 1995. Global phase velocity maps of Love and Rayleigh waves between 40 and 150 seconds. Geophys. J. Int. 122, 675–690.

Tromp, J., Tape, C., Liu, Q., 2005. Seismic tomography, adjoint methods, time reversal and banana-doughnut kernels. Geophys. J. Int. 160, 195–216.

Tsumura, N., Matsumoto, S., Horiuchi, S., Hasegawa, A., 2000. Three-dimensional attenuation structure beneath the northeastern Japan arc estimated from spectra of small earthquakes. Tectonophysics 319, 241–260.

Um, J., Thurber, C., 1987. A fast algorithm for two-point seismic ray tracing. Bull. Seismol. Soc. Am. 77, 972–986.

van der Hilst, R., Kennett, B., Christie, D., Grant, J., 1994. Project Skippy explores the mantle and lithosphere beneath Australia. Eos, Trans. AGU 75, 177, 180, 181.

van der Hilst, R.D., Kennett, B.L.N., Shibutani, T., 1998. Upper mantle structure beneath Australia from portable array deployments. In: Braun, Dooley, Goleby, van der Hilst, Klootwijk (Eds.), Structure and Evolution of the Australian Continent. Geodynamics Series, vol. 26. Am. Geophys. Union, pp. 39–58.

van der Hilst, R.D., Widiyantoro, S., Engdahl, E.R., 1997. Evidence for deep mantle circulation from global tomography. Nature 386, 578–584.

van der Lee, S., Nolet, G., 1997. Upper mantle *S* velocity structure of North America. J. Geophys. Res. 102, 22815–22838.

van Heijst, H.J., Woodhouse, J., 1997. Measuring surface-wave overtone phase velocities using a mode-branch stripping technique. Geophys. J. Int. 131, 209–230.

van Trier, J., Symes, W.W., 1991. Upwind finite-difference calculation of traveltimes. Geophysics 56, 812–821.

VanDecar, J.C., James, D.E., Assumpção, M., 1995. Seismic evidence for a fossil mantle plume beneath South America and implications for plate driving forces. Nature 378, 25–31.

VanDecar, J.C., Snieder, R., 1994. Obtaining smooth solutions to large, linear, inverse problems. Geophysics 59, 818–829.

Vasco, D.W., Johnson, L.R., 1998. Whole earth structure estimated from seismic arrival times. J. Geophys. Res. 103, 2633–2671.

Vasco, D.W., Mayer, E.L., 1993. Wavepath traveltime tomography. Geophys. J. Int. 115, 1055–1069.

Vasco, D.W., Peterson, J.E., Majer, E.L., 1996. Nonuniqueness in travel-time tomography: ensemble inference and cluster analysis. Geophysics 61, 1209–1227.

Velis, D.R., Ulrych, T.J., 1996. Simulated annealing two-point ray tracing. Geophys. Res. Lett. 23, 201–204.

Velis, D.R., Ulrych, T.J., 2001. Simulated annealing ray tracing in complex three-dimensional media. Geophys. J. Int. 145, 447–459.

Vesnaver, A., Böhm, G., Madrussani, G., Rossi, G., Granser, H., 2000. Depth imaging and velocity calibration by 3D adaptive tomography. First Break 18, 303–312.

Vidale, J.E., 1988. Finite-difference calculations of traveltimes. Bull. Seismol. Soc. Am. 78, 2062–2076.

Vidale, J.E., 1990. Finite-difference calculations of traveltimes in three dimensions. Geophysics 55, 521–526.

Vinje, V., Åstebøl, K., Iversen, E., Gjøystdal, H., 1999. 3-D ray modelling by wavefront construction in open models. Geophys. Prospect. 64, 1912–1919.

Vinje, V., Iversen, E., Åstebøl, K., Gjøystdal, H., 1996. Estimation of multivalued arrivals in 3D models using wavefront construction. Part I. Geophys. Prospect. 44, 819–842.

Vinje, V., Iversen, E., Gjøystdal, H., 1993. Traveltime and amplitude estimation using wavefront construction. Geophysics 58, 1157–1166.

Virieux, J., 1984. SH wave propagation in heterogeneous media: velocity-stress finite difference method. Geophysics 49, 1933–1957.

Virieux, J., 1986. P-SV wave propagation in heterogeneous media: velocity-stress finite difference method. Geophysics 51, 889–901.

Virieux, J., Farra, F., 1991. Ray tracing in 3-D complex isotropic media: an analysis of the problem. Geophysics 56, 2057–2069.

Wagner, D., Koulakov, I., Rabbel, W., Luehr, B.-G., Wittwer, A., Kopp, H., Bohm, M., Asch, G., 2007. Joint inversion of active and passive seismic data in Central Java. Geophys. J. Int. 170, 923–932.

Walck, M.C., 1988. Three-dimensional V_p/V_s variations for the Coso region, California. J. Geophys. Res. 93, 2047–2052.

Waldhauser, F., Lippitsch, R., Kissling, E., Ansorge, J., 2002. High-resolution teleseismic tomography of upper-mantle structure using an *a priori* three-dimensional crustal model. Geophys. J. Int. 150, 403–414.

Wang, B., Braile, L.W., 1996. Simultaneous inversion of reflection and refraction seismic data and application to field data from the northern Rio Grande rift. Geophys. J. Int. 125, 443–458.

Wang, Y., Houseman, G.A., 1994. Inversion of reflection seismic amplitude data for interface geometry. Geophys. J. Int. 117, 92–110.

Wang, Y., Houseman, G.A., 1997. Point source τ–p transform: a review and comparison of computational methods. Geophysics 62, 325–334.

Wang, Y., Pratt, R.G., 1997. Sensitivities of seismic traveltimes and amplitudes in reflection tomography. Geophys. J. Int. 131, 618–642.

Wang, Y., White, R.E., Pratt, R.G., 2000. Seismic amplitude inversion for interface geometry: practical approach for application. Geophys. J. Int. 142, 162–172.

Wang, Z., Dahlen, F.A., 1995. Spherical-spline parameterization of three-dimensional earth models spherical-spline parameterization of three-dimensional earth models. Geophys. Res. Lett. 22, 3099–3102.

Wang, Z., Tromp, J., Ekström, G., 1998. Global and regional surface-wave inversions: a spherical-spline parameterization. Geophys. Res. Lett. 25, 207–210.

Wapenaar, K., Fokkema, J., 2006. Green's function representations for seismic interferometry. Geophysics 71, 133–146.

Wapenaar, K., Fokkema, J., Snieder, R., 2005. Retrieving the Green's function in an open system by cross correlation: a comparison of approaches. J. Acoust. Soc. Am. 118, 2783–2786.

Warren, L.M., Shearer, P.M., 2002. Mapping lateral variation in upper mantle attenuation ny stacking P and PP spectra. J. Geophys. Res. 107, doi:10.1029/2001JB001195.

Weeraratne, D.S., Forsyth, D.W., Fischer, K.M., Nyblade, A.A., 2003. Evidence for an upper mantle plume beneath the Tanza-nian craton from Rayleigh wave tomography. J. Geophys. Res. 108, doi:10.1029/2001JB001225.

White, D.J., 1989. Two-dimensional seismic refraction tomography. Geo-phys. J. 97, 223–245.

Widiyantoro, S., Gorbatov, A., Kennett, B.L.N., Fukao, Y., 2002. Improving global shear wave traveltime tomography using three-dimensional ray tracing and iterative inversion. Geophys. J. Int. 141, 747–758.

Widiyantoro, S., van der Hilst, R., 1997. Mantle structure beneath Indonesia inferred from high-resolution tomographic imaging. Geophys. J. Int. 130, 167–182.

Wiggins, S.M., Dorman, L.M., Cornuelle, B.D., Hildebrand, J.A., 1996. Hess deep rift valley structure from seismic tomography. J. Geophys. Res. 101, 22335–22353.

Williams, E., 1989. Summary and synthesis. In: Burrett, C.F., Martin, E.L. (Eds.), Geology and Mineral Resources of Tasmania, vol. 15 (Special Publication). Geological Society of Australia, pp. 468–499.

Williamson, P.R., 1990. Tomographic inversion in reflection seismology. Geophys. J. Int. 100, 255–274.

Xu, S., Lambaré, G., 2004. Fast migration/inversion with multivalued ray-fields. Part 1. Method, validation test, and application in 2D to Marmousi. Geophysics 69, 1311–1319.

Xu, S., Lambaré, G., Calandra, H., 2004. Fast migration/inversion with multivalued rayfields. Part 1. Applications to the 3d seg/eage salt model. Geophysics 69, 1320–1328.

Yang, T., Grand, S.P., Wilson, D., Guzman-Speziale, M., Gomez-Gonzalez, J., Dominguez-Reyes, T., Ni, J., 2009. Seismic structure beneath the Rivera subduction zone from finite-frequency seismic tomography. J. Geo-phys. Res. 114, doi:10.1029/2008JB005830.

Yang, Y., Ritzwoller, M.H., Levshin, A.L., Shapiro, N.M., 2007. Ambient noise Rayleigh wave tomography across Europe. Geophys. J. Int. 168, 259–274.

Yao, H., van der Hilst, R.D., de Hoop, M.V., 2006. Surface-wave array tomography in SE Tibet from ambient seismic noise and two-station analysis. I. Phase velocity maps. Geophys. J. Int. 166, 732–744.

Yao, Z.S., Roberts, R.G., Tryggvason, A., 1999. Calculating resolution and covariance matrices for seismic tomography with the LSQR method. Geophys. J. Int. 138, 886–894.

Yomogida, K., 1992. Fresnel zone inversion for lateral heterogeneities in the earth. Pure Appl. Geophys. 138, 391–406.

Yoshizawa, K., Kennett, B.L.N., 2004. Multimode surface wave tomography for the Australian region using a three-stage approach incorporating finite frequency effects. J. Geophys. Res. 109, doi:10.1029/2002JB002254.

Zelt, B.C., Ellis, R.M., Clowes, R.M., Hole, J.A., 1996. Inversion of three-dimensional wide-angle seimic data from the southwestern Canadia Cordillera. J. Geophys. Res. 286, 209–221.

Zelt, B.C., Ellis, R.M., Zelt, C.A., Hyndman, R.D., Lowe, C., Spence, G.D., Fisher, M.A., 2001. Three-dimensional crustal velocity structure beneath the Strait of Georgia, British Columbia. Geophys. J. Int. 144, 695–712.

Zelt, C.A., 1999. Modelling strategies and model assessment for wide-angle seismic traveltime data. Geophys. J. Int. 139, 183–204.

Zelt, C.A., Barton, P.J., 1998. Three-dimensional seismic refraction tomography: a comparison of two methods applied to data from the Faeroe Basin. J. Geophys. Res. 103, 7187–7210.

Zelt, C.A., Ellis, R.M., Zelt, B.C., 2006. 3-D structure across the Tintina strike-slip fault, northern Canadian Cordillera, from seismic refraction and reflection tomography. Geophys. J. Int. 167, 1292–1308.

Zelt, C.A., Smith, R.B., 1992. Seismic traveltime inversion for 2-D crustal velocity structure. Geophys. J. Int. 108, 16–34.

Zelt, C.A., White, D.J., 1995. Crustal structure and tectonics of the southeastern Canadian Cordillera. J. Geophys. Res. 100, 24255–24273.

Zhang, H., Liu, Y., Thurber, C., Roecker, S., 2007. Three-dimensional shear-wave splitting tomography in the Parkfield, California, region. Geophys. Res. Lett. 34, doi:10.1029/2007GL031951.

Zhang, H., Thurber, C., 2005. Adaptive mesh seismic tomography based on tetrahedral and Voronoi diagrams: application to Parkfield, California. J. Geophys. Res. 110, doi:10.1029/2004JB003186.

Zhang, H., Thurber, C.H., 2007. Estimating the model resolution matrix for large seismic tomography problems based on Lanczos bidiagonalization with partial reorthogonalization. Geophys. J. Int. 170, 337–345.

Zhang, H.Z., Thurber, C.H., 2003. Double-difference tomography: the method and its application to the Hay ward fault, California. Bull. Seismol. Soc. Am. 93, 1875–1889.

Zhang, J., McMechan, G.A., 1995. Estimation of resolution and covariance for large matrix inversions. Geophys. J. Int. 121, 409–426.

Zhang, J., Toksöz, M.N., 1998. Nonlinear refraction traveltime tomography. Geophysics 63, 1726–1737.

Zhang, Z., Shen, Y., 2008. Cross-dependence of finite-frequency compressional waveforms to shear seismic wave speeds. Geophys. J. Int. 174, 941–948.

Zhao, D., 2004. Global tomographic images of mantle plumes and subducting slabs: Insight into deep Earth dynamics. Phys. Earth Planet. Inter. 146, 3–34.

Zhao, A., Zhongjie, Z., Teng, J., 2004. Minimum travel time tree algorithm for seismic ray tracing: improvement in efficiency. J. Geophys. Eng. 1, 245–251.

Zhao, D., Hasegawa, A., Horiuchi, S., 1992. Tomographic imaging of P and S wave velocity structure beneath Northeastern Japan. J. Geophys. Res. 97, 19,909–19,928.

Zhao, D., Hasegawa, A., Kanamori, H., 1994. Deep structure of Japan sub-duction zone as derived from local, regional, and teleseismic events. J. Geophys. Res. 99, 22313–22329.

Zhao, L., Jordan, T.H., Chapman, C.H., 2000. Three-dimensional Fréchet differential kernels for seismic delay times. Geophys. J. Int. 141, 558–576.

Zheng, S., Sun, X., Song, X., Yang, Y., Ritzwoller, M.H., 2008. Surface wave tomography of China from ambient seismic noise cross correlation. Geochem. Geophys. Geosyst. 9, doi:10.1029/2008GC001981.

Zhou, B., Greenhalgh, S.A., 2005. 'Shortest path' ray tracing for most general 2D/3D anisotropic media. J. Geophys. Eng. 2, 54–63.

Zhou, Y., Dahlen, F.A., Nolet, G., 2004. 3-D sensitivity kernels for surface-wave observables. Geophys. J. Int. 158, 142–168.

Zhu, H., Ebel, J.E., 1994. Tomographic inversion for the seismic velocity structure beneath northern New England using seismic refraction data. J. Geophys. Res. 99, 15,331–15,357.

Zielhuis, A., van der Hilst, R.D., 1996. Upper-mantle shear velocity beneath eastern Australia from inversion of waveforms from SKIPPY portable arrays. Geophys. J. Int. 127, 1–16.